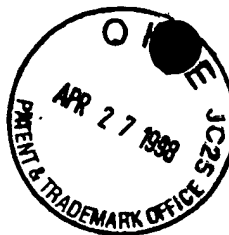


SEQUENCE LISTING



08/836,734

(1) GENERAL INFORMATION:

(i) APPLICANT: BECKMANN, JACQUES  
RICHARD, ISABELLE

(ii) TITLE OF INVENTION: LGMD GENE CODING FOR A CALCIUM DEPENDENT  
PROTEASE

(iii) NUMBER OF SEQUENCES: 67

(iv) CORRESPONDENCE ADDRESS:

(A) ADDRESSEE: NIXON & VANDERHYE P.C.  
(B) STREET: 1100 NORTH GLEBE ROAD  
(C) CITY: ARLINGTON  
(D) STATE: VIRGINIA  
(E) COUNTRY: U.S.A.  
(F) ZIP: 22201-4714

(v) COMPUTER READABLE FORM:

(A) MEDIUM TYPE: Floppy disk  
(B) COMPUTER: IBM PC compatible  
(C) OPERATING SYSTEM: PC-DOS/MS-DOS  
(D) SOFTWARE: PatentIn Release #1.0, Version #1.30

(vi) CURRENT APPLICATION DATA:

(A) APPLICATION NUMBER: US 08/836,734  
(B) FILING DATE: 02-JUL-1997  
(C) CLASSIFICATION:

(vii) PRIOR APPLICATION DATA:

(A) APPLICATION NUMBER: EP 94402668.1  
(B) FILING DATE: 22-NOV-1994

(viii) ATTORNEY/AGENT INFORMATION:

(A) NAME: WILSON, MARY J.  
(B) REGISTRATION NUMBER: 32,955  
(C) REFERENCE/DOCKET NUMBER: 960-29

(ix) TELECOMMUNICATION INFORMATION:

(A) TELEPHONE: (703) 816-4000  
(B) TELEFAX: (703) 816-4100

(2) INFORMATION FOR SEQ ID NO: 1:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 3018 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: double  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: genomic DNA

(ix) FEATURE:

(A) NAME/KEY: misc\_feature  
(D) OTHER INFORMATION: /label= FIGURE 8a

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1:

attorney: Segheto

Sub  
C#

TGATAGGTGC TTGTAAACTG TGCTTAACGA AAACATACCG TGTGCTGTAG GGA CT TAACT  
CTTGTTTATA TCAGTTAGCC TGGTTTCGCT AACAGTACAT CATTTTGCTT AAAGTACAG  
CTTACGAGAA CCTATCGATG ATGTTAAGTG AGGATTTTCT CTGCTCAGGT GCACTTTTTT  
TTTTTTTTAA GACGGAGTCT CTTTCTGTCA CCTGGGCTGG AGTGCAGTGG CGTGATCTGG  
GTTCACTACA ACCTCTCCCT CCTGGGTTCA AGCAATTCTT CTGTCTCAGC CTCCCAAGTA  
GCTGGGATTA CAGGCACCCC CCGCCACACC CGGCTTATTT TTGTATTTT AGTAGAGACA  
GGGTTTCACT ATTGTTGTCC ATGCTGGTCT CGAACTCGTG ACCTCATGTG ATCCACCCGC  
CTCGGCCTCC CAAAGTGCAG AGATTAGAGA CGTGATCCAC ATGGCCCAGC AGGACCACTT  
TTTAGCAGAT TCAGTCCCAG TGTTCATTTT GTGGATGGGG AGAGACAAGA GGTGCAAGGT  
CAAGTGTGCA GGTAGAGACA GGGATTTTCT CAAATGAGGA CTCTGCTGAG TAGCATTTTC  
CATGCAGACA TTTCCAATGA GCGCTGACCC AAGAACATTC TAAAAAGATA CCAAATCTAA  
CATTGAATAA TGTTC TGATA TCCTAAAATT TTAGGACTAA AAATCATGTT CTCTAAAATT  
CACAGAATAT TTTTGTAGAA TTCAGTACCT CCCGTTACACC CTAAGTAGCT TTTTGTCAAT  
ATTGTTTTCC ATTCATTTGA TGGCCAGTAG TTGGGTGGTC TGTATAACTG CCTACTCAAT  
AACATGTCAG CAGTTCTCAG CTTCTTTCCA GTGTTACCT TACTCAGATA CTCCCTTTTC  
ATTTTCTGGC AACACCAGCA CTTCTATGGCA ACAGAAATGT CCCTAGCCAG GTTCTCTCTC  
TACCATGCAG TCTCTCTTGC TCTCATACTC ACAGTGTTTC TTCACATCTA TTTTGTAGTT  
TCCTGGCTCA AGCATCTTCA GGGCACTGAA ACACAACCCCT CACTCTCTTT CTCTCTCCCT  
CTGGCATGCA TGCTGCTGGT AGGAGACCCC CAAGTCAACA TTGCTTCAGA AATCCTTTAG  
CACTCATTTT TCAGGAGAAC TTATGGCTTC AGAATCACAG CTCGGTTTTT AAGATGGACA  
TAACCTGTCC GACCTTCTGA TGGGCTTTCA ACTTTGAACT GGATGTGGAC ACTTTTCTCT  
CAGATGACAG AATTACTCCA ACTTCCCCTT TGCAGTTGCT TCCTTTCCTT GAAGGTAGCT  
GTATCTTATT TTCTTTAAAA AGCTTTTTTCT TCCAAAGCCA CTTGCCATGC CGACCGTCAT  
TAGCGCATCT GTGGCTCCAA GGACAGCGGC TGAGCCCCGG TCCCAGGGC CAGTTCCTCA  
CCCGGCCAG AGCAAGGCCA CTGAGGCTGG GGGTGAAAC CCAAGTGGCA TCTATTACAG  
CATCATCAGC CGCAATTTTC CTATTATCGG AGTGAAAGAG AAGACATTCG AGCAACTTCA  
CAAGAAATGT CTAGAAAAGA AAGTTCTTTA TGTGGACCCT GAGTTCACAC CGGATGAGAC  
CTCTCTCTTT TATAGCCAGA AGTTCCCCAT CCAGTTCGTC TGAAGAGAC TCCGGTGAGT  
AGCTTCCTGC TTGCTGGCTG GGTTCCCCC CCACGGAGGA GTCCTCTCAC TCAGCACCTC  
CGGCAGCTCA GCTGTGCACA TGGGCACTGG GGAAGGATC CTGGCAGCAG CTCTGCTGGG

Sub  
C

sub  
C1

CTCTGTCTTT AAGTGTGAAG CAGGGAGGAG AGGAACAGGT CTCAGATATT TCACCAAATC  
TCAGCAAAAT CCAGAGGGAG AGCGCAGGAG GTGGGGTGAT TCTTATGCTC TGGCTCTTTC  
TCTCTGAAAA AAAAAAAAAA ATCTTGCTTT TTATAAAAGT GGGTGGAAC TCACTTAATT  
CATCCTGTAA AAATAAATAT TCCTTTCTCA GAACAAATTC CAGACAGCCC AGATGTACCT  
GTTTCGTTTTA ATATTATTCA TCTTGGTAAG ATTATTTTCTG TTTCTCTGGC TAAAATCATG  
ATGTTATTCT TCTTTAATTT ACCAATGGCC ATTCTTTCTG AAACACAGAA ACCCTAGAAA  
GAGAAGAGTC ATAGGCAAGG AATTTTTTTC ATGCATAAAA TGTTGGGGTT AAAGAGAGAG  
AGACCTAGCA ATCGCTTTGG TCCACCTACC TCACCTCATA AGTGAGGAGT CAAGGCACAC  
TAGAGTGAAA TATATCTAGT GGGCACATGA CAGAGCCCGG ATTAAAACTT TGTTTTAGGA  
AACTCTCCCA GCCTCTGGGT TTCATTTACA GTGATCGCCA GGAGGGAAAT CACATTCCCC  
TGGCTCACCT CTCTGATCAT CCCTCCAGTG TGACTCTTGT TCTTAATTCG AGAAATATTT  
ATTGAGCATC TACTAGTGCC AGCACTGGGC AAGCACTGG GGGGACAGCA GTGAGTAAGA  
AAGACCAAAA TTCCAGCTGT CTTGGAACCT AGGCTCCTGA AGGGAAGATG GGCATTGAAC  
AAGAGTGACA TTGTCAGGAG ACGATGTTCT GGGTGCCACA GGATCATGTG GCAAGGAGAG  
CTAACCTGGT CCAGGGAGAC AAACCCTCTC TGAGGAAATG ATGACAAGCT GAGACCCAAT  
ACTATTGATT AGCCATGGTT TTCTTTAACC TAAGGTGGGC CAGGCATGGT GGCTCATGCC  
TATAAACCCA GCATTTTGGG AGGCCGAGGC TGGAGGATTG CTTGAGCCCA AGAGTTAGAG  
ACCAGCCTGG GCAACAGGGT GAAACCTAT CTCTTTTGTA CTAAAAATTC AAAAAATTAT  
CCAGGCATGG TGGCACATGC CTGTGGTCCT AGCTACTCAG AGGCTGAGGT GGAAGATCA  
CTTGAACCTG GGGAGTTTGA GGCAGCAGTG AGCCGAGATC ATGCCACTGC ACTCCAGGCT  
GGGTGACAGG AGTGAGAC

(2) INFORMATION FOR SEQ ID NO: 2:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 11451 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: genomic DNA

(ix) FEATURE:

- (A) NAME/KEY: misc feature
- (D) OTHER INFORMATION: /label= FIGURE 8b

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2:

Sub  
C1

GATCCACCCG CCTTGGCCTC CCAAAGTGCT GAGATTACAG GTGTGAGCCA CCACGCCAG  
CCGACACTGC CCTAACTCTC AAGTTGCATC CTTACTCGAA TAGTATGACA GTGTGGGAAG  
CAGCATGGGA CAATGTAAAA AGGAGGCATG TTTCTGGCTT CTGCTACTTA CTAGCTGTGT  
GTCTTTGCAC GAGTTTCTTA ACCTCTCTGG GCCTCAGTTT CCTTATCTGA AAAATAACAA  
TGATAGTATT CCCTTCACAG GGCCAAATGG AATACTATCA GGAACACTAC ATAATGGAAC  
TCAATAAATA ATAGCTACTG CGGCCGGGCG CGGTGGCTCA CATCTGTAAT CCCAGCACTT  
TGGGAGGCCG AGGCGGGTGG ATCACAAGGT CAAGAGATGG AGACCATCCT GGCCAACATG  
GTGAAACCGT ATCTCTACTA AAGATACAAA AATTAGCTGG GCATGGTGGC GCATGCCTAT  
AGTCCCAGCT ACTCGAGAGG CTGAGGCAGG AGAATCACTT GAACCCCGGA GGCAGAGGTT  
TCAGTGAGCC AAGATTGCAC CAGTGCCTG CAGCCTGGCG ACAGAGTGAG ACTCCGTCTC  
AAAAAATAC CTATCTATCT ATCTGTCTAT CTA CTGTTAT TCTTACCTGG TCATTTCTTT  
TTTGTTCAC AGGAAATTTG CGAGAATCCC CGATTTATCA TTGATGGAGC CAACAGAACT  
GACATCTGTC AAGGAGAGCT AGGTAGGAAA GTGCTCAGG TCAGATCCTG CCAGATGATC  
AAGGGGTGAT TACAAGGTGT GATCCCCTTC CAGGAGGTAA AGGGACAATC TGTGCTTGCT  
TCCAGTAACT TTTTGGAAGA TTTTTTATAA CAGTTGCTTT ATGGTCGTTT ATCTACATGC  
TGGCGATTGC TTCATTTCTT CCTACATGCC TCTTTAGCAC TCTGCCATGC ATCACAGGGG  
GTATCTGCAT CCTGTGGCCT CCTCTCCAGT ATCTCAAGGA CACTTACATA CCCCCTCAG  
CATGACAAAA GCCCTGCTTT TCACTGTATC GTCTTTCTTG GAAGACAGCT CTGTGACTGT  
GCACCAAGCA TGCCCCTTGG GCATGGAGAT TCTAGATACA CACACAAAAG GCATCGCCAA  
GGAAAGCACT TGTAAGTGA ACCCTTGGTT TAAATTGGCC CAGCATAGCT CCATCTTTAA  
AAGAGTCTTT CCACAAAGAT GGCATCCGCC ATGTGGATGA GCATCCAATT TTCTCTTTGA  
TTGGTTAGCT TGACTGCTCC ATCTGATCTT CCTCTCTCTC GACCTCTTGT TCAGAAAGTA  
TTGTCTTTGG TGTGGACTAT AAGCAAGCTC TGTGAAGTAA AATTGGAGAG AACACCAACA  
GAAACAATTT AAATTTGAGG AAAAGGGGGC ACCTAAGACC AAAGGAATTT GGCTTATTTT  
ATTCCAGAAG GGGAGGCTGA GAATAAATCA GATGAATATC TGGGTTCTTG CACCTGAGGG  
AAGGCTTCCT GCAGAGCCCT GGGCATAATA ATCTGGGACC TTCAAACCAA TAACCTCTTT  
TCCAAGGAAA GACTGGCTGC TTCCAAGGAG GGTAGGGGAG AGTCGGGCTG CAGGCAGCTC  
TCAAGTCTCC CCTTGCACAC TCTCAGGTTG GCATTTTCAC TTTAACCCAT CCTCCCTTAA  
GAAGGCAGTT CTTTGTGACC AGGGTACACC CCCTATTATA TATATATATA CACACACAGA  
GAGAGAGAGA GAGAGAGAGA GAGAGAAAGA GAGCAAAGTG TTACCTCCAA CTACATACAG  
TACTCTGTCA GAAAAGAGGT TCAGAGAATA AGAAAACGTC CCGAGCTCAT TCCGTTGCCA

Sub  
C1

GCAATGTCTT ACTGCCCCCT ATAGACGGGT TCCAGGGCAG CTGCCTACCT GGCCTTCCTT  
CCAATACAAA TCATCTTGGT GGATGGTTCT CTGAGGCTCA GTCTTCGCTG AAGTCAGAAG  
AGGAATTGGA CTCACATTGC AAAGGCACAG GGCAGGGCAG ATTTCTTACA GGTCTTAGGA  
AGAACAACCC AGTTATGATC ACCTACTGCT CTGTCTCCAT TGAGGCCTAA AAAGGAAGTG  
AGTTTATACT GCAGTTGGAG GAACTGCCTG CAGCCTTGAG GAAAATGTCT AGTCACAAGG  
GAGTAAGTTA CCTGTTGATC ATATTGTCAA GGAATTCCTG TCCAATTCTC CTTCCCTGGG  
TTGACACCTC TGTAAGGTCA GATCTGGAAG TAGGAGAGTG GGCAACAAGG GAGTCCCCGT  
TCAGGGAAGT GGAGTGGCTG GCTGGGATTG GGGCTTTTTT TCCCAGGAG GAGCAGGAGT  
GCTCACGATC TGTGCCCTGT GTCTGCCTGC AGGGGACTGC TGGTTTCTCG CAGCCATTGC  
CTGCCTGACC CTGAACCAGC ACCTCTTTTT CCGAGTCAZA CCCCATGATC AAAGTTTCAT  
CGAAAAC TAC GCAGGGATCT TCCACTTCCA GGTGAGGTAA TGAGAGTGTA GTTAAGAGGG  
CCAGCGGCAG GCCACCCACC GCTGGTCTCC TGGCCTTGAC TTCCCAGAAG CTGGAGGAAA  
CTTCCCACCC ATCTACCCGC AGCGGCAACA GTCGGCATGG ACCCCCTTAA GGCTTCAAGC  
CTGGGAGGAA GCAGTTGCTT ATCTCTGGCT CCCTAATCCC TCCCCACCA CTTTCCACTA  
TGTCCCAGAA AGACAGGAAG ACATCCTGTT TACTGTGGGT CTATTTTTGT CTTTGCAGCT  
GTCTGGCTGC TTTTATTGCC TGCAGCCTT CTCAAGTAGG TCCCTAAGAT ATTAGCACTG  
TGACACCACA GGACCCTTCA GGTGTACAG GAACCCCTGT CCAGGGCTCC TGTATACTTC  
TTCCTCTCTA AGGCATGGCG GTACCAAGGC TATCACTCCT CTCTTCCAAG CCCTGGAAGA  
AGAGTCTGCT TAACCTGGGG ATCAGGCTTC TTGTTTGCCC TAGAACTGAA TCTGATGGTT  
CTAGAATCCA TCCAGCTACT GGAAATTTTC TGGGTCCCAG TCACCTTGGC ATAGAGCTGG  
TGCTAGAGCA GAACCAAACT GAATTCTACC TGTGAGGGTC TCGTAGCTTC CGGGATGCTG  
GGGAGTCAGC CTGTCTCCAG CTTCAAAGGC TCCCTCATGT CCCAGGATGA CCCACATTAT  
CAGTTCTTGC TCCCCGGGTC TTGCACCTCA GCACGGAAGG CCTCAGAAAA GGTCTGTCTC  
CAGGCTCAGA CTCCCCCTCC TGCCGCCTTG GGAACATGGC ATATTTAAAG GGTCTCAGAT  
CTAAAGGGCC TTACATACAA ATATCAGATA GATTTCTGTT CTCATTTCAA TGAGGGAGAA  
AGTGCCATTG AAAAGGAGAC TAAACCACAT TTGGCCCTTT TCAGTTCAAA CTGATTCATT  
CAAAAAAGAG CGACATCCAA ACTTGAAATG ATTGAACAAT GTTCCTGCTA CAGCTAGAAT  
AGATTCTGGG TCACTTTGTT CCTCCGTTTC AATCCTTGTT CTTAGTTTG GCATCAAGAA  
ATACCTAAAT CAGCACAGTG CCTTCACTGC ATAGTTCCCA ATCCTGGCCA CATTGAATCA  
GOTGGGGGCA CCTGAGAGTG CTGACACCCA GGCCCTGCCC CAGACCTGCT GAGCAGGAGA

Sub  
C1

ATGAAAATCT TACATCCTAA GACACTCATG GAGCACCTAC TCTACCCATT ACTGGGCTGG  
ACTCTGTGGA AGACATGAAG TATATGTAAC TCACTTCCAG CTCTCAAAAA GCACCCAGTC  
CAGTTAGAGA CAGATTTACA CACCCCAAAC ACAAAATAGG ATGAACAGGC ACCCAGATGC  
AGAGTCCAGG AAATGATGCT GCTTTGGGAT TCAAGAACCC CCTGAGGAAT GTGGAGGAAG  
GACACATTTT CTAACAGTAA TTTGAGTATG TGACTCTGTG CGTGACGCTT CTGTGCAGTT  
CTGGCGCTAT GGAGAGTGGG TGGACGTGGT TATAGATGAC TGCCTCCCAA CGTACAACAA  
TCAACTGGTT TTCACCAAGT CCAACCACCG CAATGAGTTC TGGAGTGCTC TGCTGGAGAA  
GGCTTATGCT AAGTAAGCAA CACTTTAGAA TGTGAGGTGG GGCTAGAGGT GAGAAAGTGG  
GTTGCAAAAT CCAGCCGAGA CCTCACTCAC AGGAAGAGGC ATGTGCCTCT ATACGTGCAT  
ATGTGTGGGC ATGCAAGTCC AACTGTGACC CAAAGTTAGA GATCAGTTCC AGGCAACAAC  
AGCTCTAACT AAAAACATTA AATTTAAGAG TAGAATGAA GATTTGCATA GAAGACCTTT  
AGCTTTAGCT CACCATAGCG AGTTCTTTCA TTGCACCTCC ATGGTGGCAT TGCAAGTCTT  
GGGATCAGAG CATTGTCCCA GGGTCTCGAT TGGCTCAACC TCATGTGCTT ATAGAAGATT  
TATAAAGACA TGTTGTCTCT CAACTTAAAA GCTCCACCCC AGATGATAAT AATGGATTTT  
CAAATTTTGG AACAAGGTCA CTCTGTAATG CAGGCTGGAG TGCAGTGGTG CAGTCACGGA  
TCACTGTAGA TTGACCTCCT GGGTCAAGG TGCTCCTCCC ACCTCAGCCT CCCAAGTAGC  
TGGGACTACA TGCGGGCATC ACCATGGCCC TTTTATTTTT GTATTTTTTT GTAGAGCGGG  
GTTTTCCCAT GTTGACCCAG ACTGTTCTCG AACTCTTGGG CTCATACAAT CCACCAGCCT  
TGCCCTCCCG AAGCGCTGGG ATTGCCGGTG TGAGCCACCA CACCGGCAGC TGCTAATGGC  
TTTAATGCAG CCCTTCCTCA ACGTTCAGGA TGTAGTGGA AGAGCTCTCA GGAAGTGGGG  
ATAGCTGGGT TTCAATCCCA GTGCTTCTGG CTCTCTGTGG TCTTGGGTGG GTCACTTAGC  
CTCTTGAGCT CAGTTTCTTC ATTATGAAGA AAGGGAATCA TTGTTTCCAT CCCATGAGCT  
CATAGGGTTA ATGTGGAATT GATGAAAGAA CATCACAGCA TCCAAGAGGT AAAGTTCTGG  
TGGCAGTGGT ACCTGGGTTT TGTTCCCTGG AACTCTGTGA CCCCAAATTG GTCTTCATCC  
TCTCTCTAAG GCTCCATGGT TCCTACGAAG CTCTGAAAGG TGGGAACACC ACAGAGGCCA  
TGGAGGACTT CACAGGAGGG GTGGCAGAGT TTTTGTAGAT CAGGGATGCT CCTAGTGACA  
TGTACAAGAT CATGAAGAAA GCCATCGAGA GAGGCTCCCT CATGGGCTGC TCCATTGATG  
TAAGTCTGGG GTGTGGGGCA CAGGGTGGGG AGCTCCAAGT GTCAGGAAGC CTTTTACCCA  
ATGAAGGGCA GCATAGAGCT TTTGTGTGGG ACAGAGCGAA TGTTTTGTTT GAGGAAGCAG  
GAACTGGCTC TCAACTTTGA GGAAGTGGAA TTTCTCAAGG GAGAACAGTT CTTCCGGATT  
TTCAATAAAG AACTGGTCA AGGACATTTT AAGCCCTGGA ATGTCAGTGG AAATCAGTCC

AGAGGCCTGT GTCAGTGGAG GCCTCCCTTG CTGGTGCTCC TCAGTCTCAG CACGCTCCCA  
TTAAGCTGGC CACGTACTTG GCTGTGGACC TGAGCCCACC ATTTCCCTAA GAAAGCCTCC  
CAGTCACTGG GCTTTCACCA CACCTCCCCG CTTGAGACGT GGGCTTTGTG TTGTTACCTG  
GGAGAAGCTA AGCCTGCAGC ACCTTTCAGT GCAAAGAAAT GCTGTGAACT GAGACAGGAG  
CCAAGGGTAG GGAGATGGCC GCCCATGGCC AGGCCTCCTT CAGGGGGCAT GCCTTCCCTG  
AGGGCTGCTC AGTATATTGA TATGATAATC TTAGTGGTTT CCATTGGGGA GGATGGGGCT  
GAAGCTGAAT TCCTGCCCCT TCTTCTCCCA ACACGCCCAA TGGACAGCTT GGAAGGTCAG  
TTAGCACACA ACACCATGGA TGAACCTTTT TTCTGTATCA CTTTCTCCG TCTTCTCTCC  
ATTCGTGCTC TGTTGATCTC TCCTCTCTCC CTTTGTCTGT CCCATCTCTT TCTCCTCTCT  
CCTTCCCTTT CCACCCTTCT GTGTTTGTTT TCTCCCTCC CTGTGTTGTT CCCTACATTC  
TCCATCGGGC CTCAGGATGG CACGAACATG ACCTATGGAA CCTCTCCTTC TGGTCTGAAC  
ATGGGGGAGT TGATTGCACG GATGGTAAGG AATATGGATA ACTCACTGCT CCAGGACTCA  
GACCTCGACC CCAGAGGCTC AGATGAAAGA CCGACCCGGG TGTGTACACC TCCGATTATC  
AGAAGTGACC ATCCCTCCAA CCCACATGAC CCCGCCCTAT TAGTGTGAGA CTCCCCTCAG  
CAGCCAGGGC CTTACCCACA CACCCCCACC TGGCACCTCC CAAGGGTCTG GGTGAAATA  
ACTTGCTCAG CCAAGGCTCC TGAAGAGGGT GCAAGAACCA GGATTTTGGA GGAATCTCT  
GCTGGAGTTT CTGCATATTC CATEGTCCAG GCAGTTCCTC TCATAACGAA CTATCAGACA  
GAAATACTTG TAAAGATACT TCATTTATTT TGAAATATTT TTCCTCTTCT AATGTATTCA  
TTTATTCATT CAACACTTAT TTTTGAGCTC CTAATATGTT CCAGGCACTC CTCTAGCAAA  
CAAAGCAAAT TCTCTCCTCT TTTTCAATAT TTGTGGAAAA AGCAAGGTCT CCCTCTTGTA  
GAGTTTATAT TCTAGTATTT TCATAAGTTA TACCTGCTCA CTGGAGAATA CTGAGCCATA  
CAGAAAAACA CAGAGGAAAA TTCACTTAT ATTTTCCCC ATGTAAAGAT AACCACTCTT  
AACATCTAGT ATATGTTCTT CCAGGATTTT TCTATGCACA CACTGAATCT GTATTTTAT  
TTTTAAAATG TTATCATATT GTATGTACCT CTTTGCAGCC TGCTTTTTTC AGTTAGTTTT  
TTTGGTTTTT TGGTTTTTTT TTTTTTTTGG AAACCAAGTC TTGCTCTATT CCCTAGGCTG  
GAGCACAGTT GTTGCCATCT CGGCTCACTG CAACCTCTGC CTCCAAAGTT AACTAATTC  
TCCTGCCTCA GCCTCCCGAC ATAGCTGGGA TTACAGGCAC ACACCACCAC ACATGGCTAA  
TTTTTGTATT TTTTAGTAGA GACGGGGTTT CACCATGTTG GCTGGAATGG TCTTGAAGTC  
CTGACCTCAA GTGATCCACC TGCCTCAGCC TCCCAAAGTG CTGGGATTAC AAGTGTAAGC  
CAOCACACCC GGCCTAGTTT GATATTCTTA ATGTGCCCAA AGTATTCTCC TGTAACATTT

Sub  
C1

TTTAATAGCT ACACAATATT CAAACACACA GATATGTTAT AATTTATTTA CCCAATACCC  
TATTATTGGA AAGTTGAGTT CTTTTTTTTT TTTGTTTTGT TTTGTTTTGC TACTATTCTA  
AAATGCTATA ACGAACATCC CAATAGATAC ATCTTTGTAT ACATCCATGG TGAATTCCAT  
AGGACAGATT CCCAGCAGTA GAATTGCTGG GTTGAATGAT ATGCTTAGGG TAATGACAGA  
AGAGTCATTT CAAGCAGCTT CCTAGGGTCT TAGAACTTAA GGATTAATGA GTCTTCCCGC  
CCCCTCCCAG TCTATTCAGC ATGATCTGGA TCATGAGGAC TGAGATCTGG AAGAGACTGA  
GATCTGGGAG AGGCTGAGAT ACCAAAAGCC CTGGCTCCAC CCATACCCCT CGCCCTGAAA  
ACAGCTCTAG GAATTCGCG GCCTAGCAAG GCTCCGGGAA GCTCCTTTTA AAGCTGTGAC  
GTTAGTAGGC ACATGGACCA TAGAGACCTA TCCAGGGCTC ATGGGACTTT AGTGATCCTG  
CCCTTCTCCC AAGGATCCCC CATGGCTGCA ACTTGGAAT TTCTGCAAAT GGAAGAGCTA  
CTCCTTAGGC ACGGTCATGT CTGAGCAGGG ATCTCCTCGG GCTTTCTTAG AATTCTCTCC  
CTGGGCACTG GGACTCTTGA TTTCTTGAAT ATTATGTTCC AGGTGGGTGT GGAGGAGGTG  
AGGGGATGTA AAGAAGGCTA GACTTGCCA GCGCAGTGG CTCATGCCTG TAATCCCAGC  
ACTTTGGGAG GCTGAGGCGG GTGGATCACC TGAGGTCAGG AGTTCGAGAC CAGCCTGGCT  
AACATGGTGA AACCCCGTTT CTAATAAAA TACAAAAAAT TAGCTGAGCA TGGTGGCAGG  
TGCCTGTAAT CCCAGCTACT CGGGAGGCTG AGGCAGGAGT ATCGCTGGAA CACGGGAGGC  
AGAGATTGCA GTGACCCGAG ATCGCGCCAC TGCCTCCAG CCTGGGCGAC ACAGCAAGAC  
TCTGTCTCAA AAAACAAAAA AGAAGAAAA AAAGGAAAAG CTAAGACTTA CATGTGTCAC  
TTAACCCTT TTCTCAAACC TCTTCTCTT CCAGGAATAG TCAACCCTG GATGGCTTCA  
GGGGAAGGGG GATCCTGAAG CCCAGGGCAG CCTCCAATC TACCCTTCC TCCTTTGAAG  
GATACTAAGG GGTCCAGAAA GGAGGGGCAG GACACTGTTA CCCACCCAC ATCCCAGCAT  
CCACATTGCT CTCTGATGGT CAGGACAGAG CCTTCTCAGG GAGACCAGCC TGTCTGGAGC  
TGTGTCTCTT GGCCTCTTA AAGGGCCACT GAAGGTCCGT TCGTGGTCGT GAGGCACACT  
TTCAGGGAGC AGAGTGGTCT GTGTCTTAC AGAGCCCGGA AAATGAACTA GTATGAACTT  
TGCCTCCAAG CAGCAGAACT TCTGTTCCCC CGCCCCTAAT GGGTTCTCTG GTTACTGCTC  
TACAGACAAT CATTCCGGTT CAGTATGAGA CAAGAATGGC CTGCGGGCTG GTCAGAGGTC  
ACGCCTACTC TGTCACGGGG CTGGATGAGG TAAGCCTGGT GGGGCTTGGT GGGGCAAGGG  
CACCTCTCTG GGTAACTC ATGAAGTCAG GACTTAGCTG TTGGGGCCCC TGCCCTGTCT  
GCAGACTTG CCTCCAATCA GGACATTCAG TTCAAGGTCC AAGCCACGCC TGGGAGCAGA  
GGGGCTGTG AACTGGTAG AGGTGGATCC TGCCACAGTT GGTGCACAGT TTATCTTTGC  
TTTTCGTGCT AAAGATGGCA ATTTTCCAA CATTTCCAAT GAACAAATTG AAATATCACT

13



Sub  
C1

TAACTTTGCT TTTACAAAGT TGGTTTCATG TGTTCTTGAG CTTCTGTTC TCTCGTCTTC  
AGATAGCTAC AGTTGTCTCT GGGTAGCCAC GGGGACTGGT TCCAGAAGCC CCAACAGTAA  
CAAAATCTGC AGATGCTCAA GTCCCTTCTG TAAAATGGAG TAGTATTTGC ATATAACCTA  
TGCACATCCT CCCATATACT TTAAGTCATC TCTGGATTAC TTACGATACC TAACACAATG  
GAAATGCTAT GTAAATAGTT ATTGCACTGC ATTGGGTTTT TTTGGTATTA TTTTCTGTTC  
TTGTATTATT ATTTTTTCTT TTTTGAATA TTTTGTATCC ACAATTGGTT ATATGCCAAA  
GCCATGGATA CGAGAGGCTG ACTGTTCTGT TTTGCTCCTT CTGGGACTTC TGGGTTTTCC  
TGGACCATGT CTGAGACAGG AACGTTGTAA GACCTGTTGC ACACAGTTGG GCAGGTTGTG  
CCCTGTACAG AGGGATGGGC TGAGAGGGGC AGTTGCCTGC ATCACCATT GCAGCAGACT  
GGAGGGAGTC TGCTTGTGTTG TAGTTCCTCA GTCAGCAGGG GCCTTTTGTC TTTCTTCCT  
TTCCTTTTTT TTTTTTTTTG AGACGGAGTC TCACTGTGTT GCCCAGGCTG GAGTGTAGTG  
GCACAGTCTC GGCTCACTGC AATGTCCGCC TCCTGGATT C AAGCGATTTT CCTGCCTCAG  
CCTCCTGAGT AGCTGGGATT ACAGGCGCGT GTCACCATGC CCAGCTAATT TTTGTATTTT  
TAGTAGAGAT GGGGGTTTCT CCATGTTGAT CAGGCTGGTC TCGAACTCCT GACCTCGTGA  
TCCGCCCACC TCGGCCTCTC AAAGTGCTGG GATTACAGGC GTGAGCCACC ACGCCTGGCC  
AGCAGGGGCC TTTTTCTAA TTTATATGAA GACACCTAAT TTATATGTGT TAGCAAAGCC  
CTCCTGTTTA TGCCCTACCT CCTCCCCGA AGCTCATACG GCAGGATGTT CCTGAGAAAA  
TTGCCTCTTA GAAGATAGAG AGGAGATGCC AAGCCTAAGT TAGGCAGACT CAGGAGGATA  
GGTCTGACCC ACCCCCTGCC ATTCCCCAGC ACACTTGTGA TTAATCTCCT TGGCCAGAGC  
CAGGCAGAAC ACCCTCGCGT AAGAGATTTG CCCCCAGCC CCGTCCCAGC CCTCAGCTAG  
ACAGAAGATT CCCTTTCCAG AGAGGCTGCA GAGCATGAGA GCTCTTTCTG TGTGCTTAAG  
GTCCCGTTCA AAGGTGAGAA AGTGAAGCTG GTGCGGCTGC GGAATCCGTG GGGCCAGGTG  
GAGTGGAACG GTTCTTGGAG TGATAGGTAG GTGAGGGGAC CCCACGGGAT TGGCGGTGGC  
GGGGAACAGG GTCCGGGACA AGGCTGTGTT GGGAACTGAG CCATGAGAGT ATTGAAGATG  
CTTGGTATAA ATCACCCTC AAAACCAATG ATCCGCAGAG AAGAGGGGCA CAGGTGTTGG  
CTCCAGGGAA GGGCCAGGAG TGGAAGCGGG GTGCTGGGGA CCCAGAGAGG TTGCTGACAA  
CCATTGGCTG GAAAGGAAGG ATTCCAGAAA GCGTGGGGAA GGTCCAGGCA GGAAAAGCGT  
ATGAATGCAG GGTTCGGGC TAGAGAAGTG ACTTCCTTC TTGGGGTCTT GTGTTGCCTT  
TCCTGTGAAA TGGGAACAGT ATTATTAGCA CTTACCTTGT GGGCTGATAT TGAGGAGTAA  
CTGGGACTTG TTTTGGGCA AGTGCTGAGC CATTGCTAAG ATTCCCCTTA CCCGTGCTTG

Sub  
C1

TCCCTTGTAT TAAGGCACAA GGGCCCTTTG AAAAGAATTT TACCTGCTTT ATCAATTGAA  
AGGGATTAAG ACCTTGGGGG CCAACCCAAA ATAAACATGC GAACTTATTA TTTATAGGCT  
CCATGCACAC TTCGTAAAAC CTCCATGGTC CTA CTGATTACCT CCACTCAATG  
AGAGGCAATT CATTACTGAA TGAGCCATAA GCGCCTCTTA TTTCGAGAGG GGGATGGCAG  
GACTCAGTCG AGGAGAAGGA CCGCACCCAG GCAGCCTGGG CCCCTCGGCT CCTGTACTTA  
TTTACTGCTG GGTACTTCCT AGCCCAGCAT GTAATTACTG GTTCGTTGAG TCATTGCTTT  
AGTAAATGTT TCTTGGGCAC CTACTACATA GGAGGCACAG GTCAAGGCAC TGGGGATATT  
CTTTCTACCC ACCCCCTCCC TTGATACACT GTGATTAGGG ACTGACCGAT C

(2) INFORMATION FOR SEQ ID NO: 3:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1834 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: genomic DNA

(ix) FEATURE:

- (A) NAME/KEY: misc\_feature
- (D) OTHER INFORMATION: label= FIGURE 8c

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 3:

ATTTTTTTTT TTTTTTTTGA GACGGAGTCT CACTCTGCCA CCCAGGCTGG AGTGCAATGG  
CGCGATCTTG GCTCACTGCA ACCTCCGCCT CCCGGGTTC AAGTGATTCTT CTGCCTTAGC  
CTCCTGAGTA GCTGAGACTA TAGGTGCCCC CCACCACGCC CAGCTAATTT TTGTATTTTT  
ATTAGGACGG GGTTCACCA TATTGGCCAG GCTGGTCTCG AAATCCTGAC CTTGTGATCC  
GCCCACCTCG GCCTCCCAA GTGCTGGGAT TACAGGTGTG AGCCATTGCG AGCAGCCCAG  
AACTCAATTC TTAACCTTTA AAGTATGATG AGAAGAAGGA TCAAGCCCTC ACCAGCCCAT  
TTAAGGAGTT TAGGCTCAGT CTTGAGGATG TGAGAAGTCA TTGCTATTGG GTTTCACACT  
GAGGTTAACA GGTGAAGTCA GCATTTTGGT AGTTCACAGC AGCTGCAACT CTTTGTATTT  
CTCTGATACC TCCTGTCCCA ACCTACATCA GGCCTTCCCT TCTTCCTGCT TCCTTAATTC  
CTCCATTTTC CCACCAGATG GAAGGACTGG AGCTTTGTGG ACAAAGATGA GAAGGCCCCG  
CTGCAGCACC AGGTCACTGA GGATGGAGAG TTCTGGTGAG TCCAGAACCC AGGAAGACCC  
AGAAGGGTAA GGGTGGGGAA GAGAGGGGAA ATCTCAGACC TCAGTCCCCA GCTAAGGTAA  
TCAGATTCCA GCCCTTGGGA GATCTTGGCT GTGTTCTCCT CCAGCCCAAG GCCCAGCAAG  
GATGAGGTTT TGAGAGGAGC CTTCCAGGCC ACAGGGACAA TGAGCCCAGG ACCAGGCCAA  
CATGACATGG CTCTTGCCCT CTGTGTGCCC CTCCGCCACA CACTCTATTC CAGCCACAGG

Sub  
C1

CACCCTGGCC TTAGCACAAT TCTTTTCTGA GCCTAGGAAG CTCCACTTAC CCTGATCTTC  
CAACGTCAAC CTCACCCTCT CTCAGGTTGT TTCTATTAG GCTTCAAGTC TCAGCTTAAG  
GAGAATTTTC AAGTCTCAGC TTAAGGAGAG CCCCCTAAGT TCCCCGAGGA CTGGGATTAA  
TTTATGATGC TCATCACCCCT TAAAATTGTT TGCTTAAGCC GGGCGCGGTG GCTCACGCCT  
GTAATCCCGAG CACTTTGGGA GGCCGAGGTG AACGGATCAC GAGGTCAGGA GATCGAGAAC  
ATCTTGGCTA ACACGGTGAA ACCCTGTCTG TACTAAAAAT ACACAAAAAA AGTAGCCGGG  
CGTGGCAGCG TGCGCCTGTA GTCCTAGCTG CTGGGGAGGC TGAAGCAGGA GAATCACTTG  
AACCTGGGAG GCAGAGGTTA CAGTGAGCCC AGATTGCGCC ACTGCACTCC AGCCTGGGCG  
ACAAGAGAGA CTCTGTCTTG GAAAAAAAAA AAAAAATGTG GTCTTAGTTT AATGTCAAGG  
GAAAGGTTTT GGGTGTTTTT ATTACTTTAT TTTTTATTTA AAAACTATAA TAGAGACGGG  
CCTCGCTATA TTTCTCGGGC TGGTCTCAA CTCCTGGGCT CAAGCGGTCC TCCCACCTTG  
GCCTCCCAA ATGCTGGCAT GTGGGCCTGG TCAACATATG GGACCCCAAC TCTACAAAAA  
ATTTTAAAT TAGCCAGATG TGGTGGCGTG TGCTGTAGT CCCAGCTACT TGGGAGGCTG  
AAGCAGGGGG TCACTTGAGC CCAGGAGGTT GAGGCTGCAG TGAACATGA TTGTCGTTCA  
CTTTTCTTCT GAACGTGAGA TTAAGTGTA TCAGCAATTT GGCTTAGGAT TATTTATTCA  
GAATTTTTAA CCGTCACGTT GCGGCAACC AGGT

(2) INFORMATION FOR SEQ ID NO: 4:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 14664 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: genomic DNA

(ix) FEATURE:

- (A) NAME/KEY: misc feature
- (D) OTHER INFORMATION: /label= FIGURE 8d

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 4:

AGGAGGTGGA GGTGTCAGTG AGCCAAGATC ATGCCACTGC ACTCTAGCCT GGGCAACAGA  
GCGAGACTCT GTCTCAAAAA ATACACACAC ACACACACAC ACACACACAC ACACACACAC  
ACACAATAT ATATACACAC ATATATATAC ACACACATAT ACACACACAC ACGTCTGTAT  
ATATATGTGT GTGTGTATAT ATACACACAC ACACTATTCT ATATATTCTT GTAGAGCTAT  
GTGTGTCTCC TGTGCTATTG AGCATGAGCC CTTTTTTTTT TTTTTTTTTT TTGAGACAGA  
GTCTCACTTT GTCGCCCAGG CTGGCATACA ATGGCGCAAT ATCGGCTCAC TGCAACCTCC

Sub  
C1

GCCTCCTGGG TTCAAGTGAT TCTCCTGCCT CAGCCTCCCA AGTAACTAGG ATTACAAGTG  
CCCGCCATAA TGCTCAGCTA ATTTTGTAT TTTCAGTAGA GATGGGGTTT CACCATGTTG  
GCCAAGCTGG TCTCAAACCTC CTAGCCTCAG GTGATCCACC TGCCTCAGCC TCCCAAAGTG  
CTGGGATTAC AGGCATGAGC CACAGCACCC TGGTGAGCAC TAGAGCTTAT TTCTTCTATC  
TAACTGTATT TTTGTATCCA TTAGCCACCC TCTTTTCATC CTCCCCTCTC CTTCCCTTCC  
CAGCCTCTGG TAACCACTGT CTGCTCTCTA CTTCCATGAC ATATGCTTTG TTTTAGCTCT  
CACATATGAG TGAGAGCATG CGACATTTAT CTTTCTGGCC CTGGCACATT TTTGAATCAT  
TGTTAGAAAA GATGATGGTT TGGAGTAGAT ACATCAGAAG TGACAGCGTT TGCCCTAAAA  
AGGAAAGACA GGCTCCTCTG GGACCCCTGAC CAAGTTCCTG TGAACATTTT TATTATTGTG  
CTGTGTTAGT CCTGGGGTCT TCCGTTCCCA GCCCTCCTCA CCTGCTCCCA TATGGCTCTC  
TCTCTTCTTC CAACCTCTCA GGATGTCCTA TGAGGATTTT ATCTACCATT TCACAAAGTT  
GGAGATCTGC AACCTCACGG CCGATGCTCT GCAGTCTGAC AAGCTTCAGA CCTGGACAGT  
GTCTGTGAAC GAGGGCCGCT GGGTACGGGG TTGCTCTGCC GGAGGCTGCC GCAACTTCCC  
AGGTGGGAGA TGCTCTTGAT GGGGGGAGGG TCTAAGCCGA AAAAGTTCCA GGCAGAAGAA  
GCCTAACTAG TGCTTATTAA GTCTCTCTGT TCCAGACGTC CACTATCTTA TTAAACCTTC  
CCTGTTTTAC TGAGAAGGAA ACCACCATGC TGAGAAGTTT GCAATAGGGA GCTGGGTAGC  
AACTTTGGAA GCAGGAACTT GTGGGAACAA TGCAGATGCT GCTTGGACTT ACGATGAGGT  
TATGTCCAGA TAAGCCCATC CATCTTTTGA AAATACCCTA AGTGAAAAGT GCATCCAATA  
TGCCTAACCC CCCAAACCTC ATAGCTTACC CTGGCCTACC CTCAAACATT GCTCGGAACC  
CTTGACCTTA AGCCTAAAGT TGGGCCAAAT CATCTAACTC CAAAGCCTAT TTTACAAAGA  
AAGTTGTTGT AATATCTOCA TGTAACCTAC TTAATACTTG TACCTAAAAA GTGAAAAACA  
AGAATGGTTG TACGGCTACT CGAAATCCAG TTTCTACTGA ATGTGCATCT CTTTCACATT  
GTAAAGTTAA AAAATTGTAG CCGAACCATC CTAAGTCAGG GACTGTGAGT ACTGTGTCAG  
TAACAGTAAG GGCACCTATT GAGAACCAAG TTAGCAGCTG CTGCAATAGT TCAAGTCAGA  
GATGATGAAA ACCTAGACCA AGTCAGTAGC AGCAGAGATG GAGGGGAGAC AGCAGATTTA  
GGGAGAGCAT ATTGGGTGAT GTAGGGAAGG AAGAAGAATG ATGTCAAGAT TCCAGTTGG  
GGACCTGACA ACATTGCAAC ATAAGACACA CAAGAAGATC GGGTGGGTGG CTCATGCCTA  
TAATCCAGC ACTTTGGGAG GCAGAGCCAG GAGGATCACT TGAGCCCAGG AGTTCAAGAC  
CAGCACAGGC AACATAGTGA CACCTCATCG TTACCCAAAA TAAAAAAAAA AATGAGGTGG  
GAGGATTGCT TGAGCTCGGG AGGTTGAGGC TACAATAAAC TGTGATCATG CCACTGCACT  
CCTGCCTGGG TGACAGAGTG AGACCCTGCC TCAAAAAAAAAA AAGACACACA AGAGAAAAAT

Sub  
C1

ATCAGCGTGT TGTTTGT TTTT TGGTGGAGTT AATTGTGGGG TTCTAGGGAA AGGAATT TAG  
CTTGGGACAT GGAAAGTTTG AGGTTCTGT AGAGTGTCCT AGTGAAGATT TGTAAATAGAG  
CATCGGATGC GCATATTAGA TGGCACTTGG TGATATGATA AGAACTCAAA AAATATTTGA  
GGAATAAAGG AAAGAAGAGG CCAGACGTGG TGGCTTATGC CTGTAATCCC AGCACTTTGG  
GAGGCTGAGG CAGGCGGATC ACTTGTGGTC AGGAGTTCGA GACCAGCTTG GCTAACATGG  
TGAAAACCCA TCTCTACTAA AGATACAAAA ATTAACCGGG GATGATGGTG GGTGCCTGTA  
ATCCCAGCTA CTTGGGAGGC TCAGTCAGAA GAATCGCTTG AACCAGGAG GCGGAGGCTG  
CAGTGAGCCG AGATCGCGCC ACTGCACTCT AGCCTGGGCA ACAGAGCCAG ACTCCGTCTC  
AAAAAAAAA AAGTGAGAGA GATTGAGGCT GGGATATATG GCTCAGGCAT CATGCGCGTG  
TAGGGGGCAG TTA AAAAGCA GAAGTAAGAA AGATTGCCTA GGGAGGCAGG AAGGGTGAGG  
TGAGAGGAGA AGAGGCCAG GACCAGATTC TAGTCACCAA CAGCGTTTAA GGGGCAGGTA  
AGGAAAACAA AACCATCAGC AAAGACTGAG AATGAAAGCC CAGAGAGGAA GGAAAAGCCA  
CACATACAAT CAGTACAGCT CCATCTGAAT AAAGGTAGCG CCCCCCCCCC CCCAATCAT  
TAGAGAAATG CCTGATTTCG TTTTCTGTGG ATTTTTCCTA AGAACCTAGA TGTGGGGAAT  
AGAAATAAAT GGTTCCTCT GTCTCATCCC CTCCCTGCCC TCTGAGAGGA AGCTGTGATT  
GCGTGCTCCC TTTCTGGGGG TGCAGATACT TTCTGGACCA ACCCTCAGTA CCGTCCGAAG  
CTCCTGGAGG AGGACGATCA CCCTGATGAC TCGGAGGTGA TTTGCAGCTT CCTGGTGGCC  
CTGATGCAGA AGAACCGGCG GAAGGACCGG AAGCTAGGGG CCAGTCTCTT CACCATTGCC  
TTCGCCATCT ACGAGGTGTG TAGTCCTGAT TGGCTCCAGC CCAGGAAACA TACTTTCCCA  
GAGAGGACGC TTCCAGGGGC TTCTAGAGGG GCCCTCTGCT TCCTCAATAC CAGTGACCCA  
CAGAGCTCCT GGTATCAGGA CCACTTGTGT TTGTAACAAG CAAAAAATAC CAGGGGGGGC  
ATTAGAGAGG CAGTGGAGCG GGCCTGGCAG AACAGGTGCC TGGGGGTCAG GCTTCCGCAT  
GCGGGCTCCA GTTGCTGGCA TTGCCTTCCG CAGGCTCCTC ATCCTCATTC ACATCTGAAG  
CATCTTCCTT TCTGTTTCTT CTCAAGGTTT CCAAAGAGGT ATAGCAGCAG CAGCGGCCAG  
CAGTTGTGTG CAGCACTACC CAGGGGGGCC CGAGTCTGTC TGTGGCTCGT CGAGAAGCTT  
CCTGGTGGGG TTTGTGGGCA GGA CTGTGA TAGGAGAGGG CCTTGCCTGT TGTTATTTCC  
CACTTGACAGA GCAGGTTGCC TCAGGGCATT GCATGACCCA TGACTACCAC CCCCAGGATG  
TGCACTTTCT CCCTCGCACC AGACACTGCA CGTCACACAC ATGCCTTTGC AACTCACCC  
TCCTCCACGC TTACAGCCAC ACACACAGTC ACACAGACGC GTTCTGAGGG TGGCTGCCCCG  
CTTGGGATGG AGGAATCACT TCCCTCAGAA CCCAGCCAAG TCCTCTAGGC CTCCTTGGGG

Sub  
C1

GTCCTTCCAG CCTGAGGGGC TTCGGAGCTG AGGACAGCTG TTCTGGTAAG TGTCCCTGAG  
TGTGGGGATG ACACATTTCC ATTCACCTCTG AATCACAACA GAAAAGGGAA GAGGAATTGA  
GGTAGGGAGC CTATTTAACC CTTGGGAGTC GGAAGTAGG GAGGTTGAAA CTGTGACATG  
GGTGACCAGG GAGTTGGGAA GGGACCCTTG GAGGTGGCTG TGGCAGGACA GGACGTTCTT  
CCCGAGGGGC TCATGTGCCC TGGGCTCTCC CCATCTCTCA GATGCACGGG AACAAGCAGC  
ACCTGCAGAA GGACTTCTTC CTGTACAACG CCTCCAAGGC CAGGAGCAAA ACCTACATCA  
ACATGCGGGA GGTGTCCCAG CGCTTCCGCC TGCCTCCCAG CGAGTACGTC ATCGTGCCCT  
CCACCTACGA GCCCCACCAG GAGGGGGAAT TCATCCTCCG GGTCTTCTCT GAAAAGAGGA  
ACCTCTCTGA GTGAGTGCTG GCCCAGCTTT CCCACGTGTT TCTAAAAGCT CACATGGCCC  
ACTCCAGAGG TTGAAGGCAT GAGGCAGCTA GACACGTCTC CTCCAGGGTC CTTCTGCTGC  
TCCTGAGCCA CTGGCCACAT TACCCCCATT CATTCAATTCA TCCATTCTGT GATATTTATT  
GAGCACCTAC TATGTTCCAG GCACTGTCCT AGGCACTAAG GATAGAGTAG TGAAGTAAAC  
AGAAAGAAAT CCCTGCCTTC ATGGAGCTTA ATATTCTAAC ATGAGACAAT AATGGATAGG  
AAAAACATAT GTAGCATGTT AGATTTGGAG AGGTGATATG GAGCAAAAAT AAAGTAGGGA  
AGAGGGATAG GAGGTGTTGG GGATGCTTGA AATTTTAGGT TAGCATGGCC AGGAAAGCCA  
CATCCTGTCC CTGGCCACCA CAGATGAGCT CATAGCCCCT GCCACTCTGA TCTCTGTCCT  
TGGAAGATGC ACCAGGTCCA TGGTAGGTG GCTGGGTCAT GCCTTTGGGG GGCTCTGAGC  
AATACTAACA AGAACCTGCG TGCCTGGGCT TGGCTGTCGG GGATGGTGCT GACATGGGGC  
TGGTTCCTGG GGTGGGGTG TTCCAGGGGT TCTCTAGAGG CTGGTTCTGG CTTGGCTGCC  
AGGAAGCCGT GCACCAGAGC AAACCGTCCA CGGGCCTCCT GCTTGCTTCT GGTGACACTG  
AGACCCACA TGTCTGTATT CCTCACAGGG AAGTTGAAAA TACCATCTCC GTGGATCGGC  
CAGTGGTGAG TGGTTTAGAT CTTCTGTGCG AAAAGTCCAG AGGGTCCCCT TCCCTGACCA  
TGCAGGGGAC AGATGGTGCA GGGGAGAATG GGCACCTGGCA GAGGGAATGG GAGTCTGGGC  
TGTGCTGAGC AGTCCCTCCT TGGCACTGCA AATCCTACTT TGGCATGGCC AGAAGTAATC  
GGCCTTAAGC ACCGGGGGCC ATTGAGGCAG TTCAGGGGCT GGGAAATATG GAAGAGGGTC  
CTGGAAAGGA GAAGCAATTT GAACAATCGG AGGGAACAAG GCCACAGGAA GGGATGACAA  
GAGCCGCAGC GAACACTGGA TTCTGAGACT GGATAACATT GGATTTTACA CATAGAGAAA  
AGAAAGTAAG CTGGTGCCGG ACCTGGTGTT GACACTTGGA TCCTCCACTT ACCAGCGGGG  
TGACCTGGAC AATTTCTGTA ATCCCTCTCA CTCAGTTTCC TACTCAGTAA AACGGGGATG  
ATAATGTGCC TTGCAAGGCT TTTGTGAGGC TTCATCAATG AGGTGATGTA TGTGAAGTGT  
CTGGCACAGC ATGGGCACTC AAACAGAGGT GCTTTTTTAC ACTTTACACC TTACAAGGTA

CTTTTCACAT GTGTCATCGC GATACTTGCA AGGTTGCTGA GAGGTAGATG GGGTTATAAT  
CCCTGGTGTT CAAGAAAGGA AGCAGAGGCT CAATGGGGTT GAATGACTTC TCTGAGTTCA  
CAGAGCTCAG TAAGTGGCAG GGTTTGAAC TCACATTCAG ACTCTCTGAC TCCAGACTTA  
GGTTTTTCCG CACCTCCACG CTGAGGCCAG CCCCAGGCAG TGAGAAGCCC AAAGTCCGAA  
GCACAGAGTG CTGTGTGTTG GGCTCTGTGT GTTGAGGAGT CTTGTGACTG CCTTGGGGCT  
TTGGGCTGTA GTCAGCTGAC AGTCCTTTGT GCTCTGTGGG GATGACGTAG GCCAATGGGA  
GGACAAATGC CCCTCTGAAC TGTCTTCTGG GCAGTGACAG TCATGGTCAT AATCCTGACC  
CTGAGCCAGT GCCAGGTCTC CAAGTGCCTT CTGAATGACC ACAGGCGATT GGTTTTAGTG  
GTAGGTGCGT GGGGATCTGT TCTGGTCATC TGGATGCTGG TCATCGGGTG CAGTATTGAT  
CAGGACCTGC AAACCCAAAA GCTTATGGGA GCTGCCACGT CACGTGAGTA GAGCAGGCAG  
GTGCAGGGTT TTTGATGTCC CTGCACTGAC ACAGTTGTCT GCAGTTCTCC AATTTGACAT  
TTGGGCTCCA GTGTCGAGGG TCAAACAAGG AATTTTGGGG CGTGGGCCAA ATCTGGGAAG  
ACACAGGGAG CAGGGCCCTT TGGCTCAAGC TGATAGTTGC CGCAGGGATT ACCAGGCCCA  
GGGCAGCCTG CCACAAGCTG GGGCTTTTAC CAAAGAAAAT CTCCTATGT TAAATGCTTG  
CTCAAAAATT TTTAAAAAAT ATTCTGTAAG TCAAAATCCA TTGTTAGGTC AGTTTGAGAG  
AGCCATGTTT TTGGTGTTTT AGTAACCAAT TTCATTTTTT TATTATTTAT TTATTTGTTT  
ATTTTTGAGA CGGAGTTTCA CTCTTGTCAC CCAGGCTGGA GTGCAATGGC ATGATCTCAG  
CTCACTGCAA CCTCCGCTC CCGGGTTCAA GCAATTCTCC TGCCTCAGCC TCCTGAGTAG  
CTGAGATTAC AGGTGCCAC CATCACGCCT GGATAATTTT TGTATTTTTT AGTCGAGATG  
GGGTTTCACC ATGTTGGCCA GGATAGTCCT GAACTACTGA CCTCAGATAA TCCGCCCACC  
TCAGCCTCCC AAAGTGCTGG GATTACAGGC ATGAGCCAGC ACGCCCGGCC ACCAATTTCA  
TTTTTTAAAA AAGGAAGAAA GAAAACCTTA GCCAGAAGAT CTTTTCTCTT GCCATATGCA  
GTAAGAGTAG ATTATAAAAA CAAAGTCAGA GCAGTCACTG GTGTCTGGGC ATGGAGGAGA  
AAGAAGAAAT CTCTTCTCCC TTCACCCTCC ATGCCCCTTT TTGGCTCCAT GTGATTGAGA  
TTTCTGACC CTGGAGCCCC ACCCAAGCT AAAGACCAGG ATACAGGGAA GCCACAACCA  
CTGGCGGTTT TGAGAACTTA CTTTTCACTT ATTCTGCATT TACTGTTTCC TTTTCTTATG  
CAGAAAAAGA AAAAAACCAA GGTAGGTGTG TGGGTAGAGA GCATGAAGTG TGTGTACTCA  
TGCATATGTA TGTGCATGCA TGTGAAGTGT GCATGTGTGA GCTCATATGC ATCCATGCAC  
CAGACTTGCC TCTTCTCCC CCTCCTTCTT GAGCTTCTGC TGGGGCCGAG CGTGCAGTAA  
TGACAACTAC GATTTGCTGG GGAAGGCTA CGTGCCAAGC ACTCTTTTAG GTGCTTTCCA

sub  
C1

TGATTAATTC CTTCTCACA ACAGCCCTAT GAGATTAGTA CTATAACTAT CCCCATTTC  
AGAGGGAGAA AAGGTACAGA CTTGACTAAC TTGCCAAGG CCACACAGCC AGAGAGGGG  
AGAGCCAGTA CTTAGAGCCA GGCAGTCTGG GTCCAGAGTC CGTGTCTGA ACCACAAGAG  
GCCATCATAC GCCATCAGAT TTGGTGCTAG CATTCTGGT GGTGCCTGGT GGTGATGGAT  
CCATCACAGG GGTCTCCAG GTACTGGTGC TGGCCAGAC CAGAGCTGAC ACTCCTCAGG  
CACTACCACA TTCCAGGCAC TGTGCTTGGG GTCAGTCCCT CTCTTTTTT TCCCCCAA  
TTATAACAGT ATCTACAAAG TAGGTGCTGT TATTTTTCCC CTTTCACAGG TGAGATAGAC  
TCAAAGAAGT GAACTGCCC AAGGAACAGA ACTAATGAGT GGGGAAAATG GAACTGGAAA  
CCATGTCTGT TTA CTCCAAA ACCTGTGTTT CTTGCCCTCT TTCTCTGATG CCAGCCCCCT  
ACACTTCAAG GCCTGTGTTG TCCAGACCCA CACTCGGGC TGCCAGTGTG TGCCTGGCAG  
GGATGCTCCA TGGCCACACC ATATCCATCC TACACATCCC CCTCAGACT GTGACCTCCA  
TTTGCTCTGG GATCCCCACA AGCTTCAGCT GCTTGAGCAA GACACTGCTT AGAAGGCAGA  
GCAAGCCAAG GCCTCTGGGG CCTGCTGGGA GCGAAAGCTG GGGAGCCGTT TCCACGGGTC  
TATCTGCTTG AGCTGTCTTA GATGAGCAGC ATGGAAGGGC AGTGGTGCAT GAGTCCAGGC  
GGGCTGCTTT TCTGCTCCGA GAGGCTCTGC CTGCCCAGTT GTTCTCTGCA TTGCAGCCTC  
AATCCCCACA GCCTTGCTT CCCCCGGCTT TCCCTACAGG TGCACCGCAT CCACAGTGTT  
GGCACCATGC AGCAGCCGCT CTCCGTCCTT TTCATATCCT TGTCACCTGC ACGAGCATGT  
CTTGAAAATA TCCCTTGTTT GTGTAGCATC TTAAATGTTT TTGCAGTATG ATTTTGCATT  
CAGTATCTCA TTTGATCCCC ACAAGAGCCC TATGAGGAGG GAAAGCAGAT TTTACCATTA  
AAGGATGAGT AAACTGAGGC CAGAGAGGAT ATTTTGGTT TTTTTGAGA CAGTCTCACT  
CTGTCACCCA GCCTGGAGTG CAGTGGCTTG ATCTTGGCTC ACTGCAAGCT CCACCTCCCA  
TGTTACACC ATTTTCTGC CTCAGCTCC CAAGTAGCTG GGA CTACAGG CACCCACCAC  
CACACCAGC TAATTTTTT GTATCTTTAG TAGAGATGGG GTTTCACCCA GTTAGCCAGG  
ATGGTCTTGA TCTCCTGACC TTGTGATCTG CCTGCTTCGG CCTCCTAAAG TGCTGGGATT  
ACAGGCGTGA ACCCCCCTGC CCGGCCAGAG AGGATATTTT TTAATGAGGG GCAGGGCTGG  
GATTCCAGCC CAGTGTCTG ATGGCTCACC CACTGACCAT TCCACTAATC CGTGTCTTTT  
TTCAATCTAA ACTTTCAGGG TTGTAGAGGT TCCTTTGAGG TGCCTCAGTA CTCCATGGT  
GATGTGGGGT CTGAGGGCCA AGAGCTCTGT TCTCATTAA CAGAGAAGCT TGTGTTTTTA  
AAAACCCAT GTTACTGCA GGAAATTTAA TTGGACAGTG TTTCCATCTG GAAAAA  
AGTCTACAAA ATACTGACA ATCACTGCAC TAGATCATGC TGCTTTTAGC ATTCTTAGCA  
TTTCACGTGC TGAGCTCTCA ATACTCTACC ATGAGGAGGG ATGGAGTGGG TATGAAAAGA



Sub  
C1

TAAAGAACTG AAGTCACACG GCTTGTCACT GGCAGAGATA GAGCTTGAAC CGAGGTTGAA  
GAGCTCCCGC CTATTCTTTT CCTCTTCTCA CTGGATAAAG CTGCTCCAAG AGAGGTGCTG  
CCTCAGTGTG CCTGTTTACA CTGTAATCCT CCCTTCCTTC CTGCCTCCTC CCTCCTCTCT  
CCAGCCCATC ATCTTCGTTT CGGACAGAGC AAACAGCAAC AAGGAGCTGG GTGTGGACCA  
GGAGTCAGAG GAGGGCAAAG GCAAAACAAG CCCTGATAAG CAAAAGCAGT CCCCACAGGT  
GTCTGGGCAT GTGGCATGGG TGGGGTGGCC AGCACGCTAC AGGGGCTTCC TATGCGCTTG  
GGATACACAG GGGCTGGAGG CTTCCAGGA GTTGTCTTG AACATCTGGA GGTTTGAATT  
TGTCCCACTG ACCTTTTCTT TCAGCAAGTT CCCCTGAAAT TTGGGCTGCT GCTTGGGTGA  
ATATCCCAGG ATGGGGGTTT CATTCTAGGA GTGGACTGGC AGGCTGAGCC TCCCATGGAG  
CTGATCCAGC CAGGATACAG AGAAGGGGAG GCAAAGGCTG AGACAGAACC AGCTTGAGAG  
CGGAGGCGCA ACTCTTGTCT CCTGGTGGCC TTGAGCATTT CACAATAGGG GGATAAAGGA  
TAGGAGCAGA AAAGTGGGGC TGACTTCAGA AATGGGGTCC TCTAGAGCTC ACGGGAGGGT  
GTTAGATTGG AGTGGGAGCT TAGTGGAGGT GAGCCTTAGA GGCAAAAGTC TCCAGACCAA  
TCCAGGCCCC CTCTTCTATC CGGGGGGCC TCTTCTATCC AGGGCCCCCTC TTCTGTCTGG  
GAGCCCCTCT TCTATCTGGG GCCTCATGCA GTGGGGCCTA GGGGAGGTTC TCTGAGGACT  
TGGCCTTGAT GACAGGGTGG CTGGAGGAAT CAGAACGGTC AGACCTTCTT TGACCTGCGG  
GCACCTTTAG TTGGAATGCT CAGGCCTGGG ATGGTGGAGG GGGCTCTTGC AGGTGGGGAC  
TGGGGTGGCG GGGAGGAGGC TGTATGGCCG CCATATCTCC TTTGGCTGGG GCGCTCAGGG  
CTGGAGAGGT GTGAAGAGTC CCTGAGGCCT CGATGCATCT CACTCCAGCT CACCAGGTCT  
GCATTTGCCC GTCCCCAGCT CCTGCTGCCA CCCCCGCCG TTTAGGCAC TTGGCTCCCT  
TGGCCCAGAG GAGCTTGCCT CACAGGCCTG TGCACCTCTG ACCCCTGTGA ACCAGTTTTC  
CTTTGTGCCT CCACAGCCAC AGCCTGGCAA CTCTGATCAG GAAAGTGAGG AACAGCAACA  
ATTCCGGAAC ATTTTCAAGC AGATAGCAGG AGATGTGAGT ACCTCCAAGC CCAGGACGCC  
CACAGGTGCT TCCTTCTCTC CTGGATTAAC TGCTCAGATT ACCAATTATT TCATTATTGT  
TTGGTAGAGG TCACTTTGGA CTTCGGTGGG GCCAGGGGAT GTGTGCGTAG CACACAAATC  
CACAAGCCCT TGAGTTTTGG ACTGCCACGT CTGCTGGGGG GCTCAGAGGC CTTTTGCTC  
TGAGCTGCCC ACGGTGGTCC TGATAGCTGA GGTGCAGTAT CTGGCCCCCT GTCTTCCTCA  
GAAAGCCCC AGCTTCCCAT GACATAATAG CACCGACAGG GATTTTACAA ACACAGCCAG  
GTGGAATTTG TTTTGCAAAG TGTCCGCGCC AGGAGCTGCT GTACTCCTGA ACCATGACCC  
TCCTCTCCCT TCCTCTCAG GACATGGAGA TCTGTGCAGA TGAGCTCAAG AAGGTCCTTA

ACACAGTCGT GAACAAACGT GAGTTGCTCA AACCAAATGG GGGTGGGGTG GGTGGGGAGT  
CCCGTTGTCT CAAAGCAGCT CCTCACTCTT CTCCATCCCC CCAGACAAGG ACCTGAAGAC  
ACACGGGTTC ACACTGGAGT CCTGCCGTAG CATGATTGCG CTCATGGATG TATCCTTCCT  
GCCGCCCTT CCCGACCCTC TGTCATCAGC CCACGGGGGC CAAGGCAACA TACAGGGTGC  
CCAGTCAGGC AAAGGGCCCT AATTTGTGCC CAGGGAACT TAAGGAGACC CTGATTGAGA  
ACATCTTGGA TACTCGTCTG AAAGGGGTG TTAGAGGCGG AAGGGGAGGA TGTGGGGTG  
TAACTGCCCT AACCCCTGTG CTTCTCTCAG GCCTGGGATC CTGCCCAAGC AAAAGTGGTC  
CTTAGGAGAG CGGCTCCTGG GTTACAGAGT AGGCGCAATC TCTGACTGGT GGTGGAGTGG  
AGGGGAGGGT TAAATAGTAC AACAGGGCAG TGGGTAGGAG AGCCCGGAGT CTCCTAGACC  
CTCCCTCCAA ATCCAGGGGG ATTTTGTCTG GTGCTGTGTA GCCCTGACCT CCCTCCTCCA  
GACAGATGGC TCTGAAAAGC TCAACCTGCA GGAGTTCAC CACCTCTGGA ACAAGATTAA  
GGCCTGGCAG GTGGGAAGAG AAAATGAAGC GTGGGAGTCA AGAATGGGGT TGATTTGGAG  
ATTGAGTGTG TGACCTCCAT CCTCAAATTT TCTATTGCCA GAAAATTTTC AAACACTATG  
ACACAGACCA GTCCGGCACC ATCAACAGCT ACGAGATGCG AAATGCAGTC AACGACGCAG  
GTGCTGAGAA GGAAGGGGTG TCAGGGATGT GGACCCGAGA CGGTGGGAGC AGGAATGGGA  
GGGGACTAGC TACTAGGGCC CCACTAGAGA AGGAGAGGGA AAGGGCTTCT CACTTTCCCT  
TCCCAGGTCA CAGAGTGTCC GAGAGGCAGG GAAAATAGAA GACAGGCCCA AGGCCTCCAG  
CTCCACGTCC ACCTCTAACA TGGTCCCCTC CACAGGATTC CACCTCAACA ACCAGCTCTA  
TGACATCATT ACCATGCGGT ACGCAGACAA ACACATGAAC ATCGACTTTG ACAGTTTCAT  
CTGCTGCTTC GTTAGGCTGG AGGGCATGTT CAGTAAGTGG GAGAGGGGGG CTGCCCTCTG  
CTCTCTTGCA GGGGCAGTTG TGGCAACAGG CATCTCACCT GATAATCTCC AGTCTGCTCC  
ATCCAGGCTG AACGAGGGCC AATGACCTCT TTAGGCCCAG AATGGGATGG CAAAGGGAGG  
GTTACTGGTG ATTCTCTGCC TGCACATCTT TGTGCTGATG AGGGACAGCA CTGGGCACAC  
GGTCTCTGA GGGGAAGTTA CAGTAGTAGA GGCGGAGTGC GCCTGTAACT GGCCTCTGGC  
CTGTGCATT TTTACAGGA GCTTCTCATG CATTTGACAA GGATGGAGAT GGTATCATCA  
AGCTCAAGCT TCTGGAGGTA AAGCATAGGC ACAGCACATT CCCCCTACAC ATTAAACTC  
AAGGTGGAGG GGTCAACGGG GCGGACTGGA CCCAGGGTGT GCTCCTCATT TCCACACAGT  
GGTGGAGGGA AGGGATAGGA ACAGAACATG GAGGGAGGCT CAGCAGGCTC CCAGGACACA  
TGCCTTGAG GCCAAAAGG ACCTCTGCTC CCCAGTCAC TTGATGCGGG AAAACATGCA  
CCTTCTTAGG GAAGATCTAG GAGAAAGGAA ACAGTAAGCC ACTGCTTCTT GGAAAATCTT  
CTGGGGGTCT GACCTGCTGG GACTGTTCCC TTTCTCTTG CCCCCTAAGA TTCCTAGGGC

Sub  
C1

Sub  
C1

GGGGGGGGGG GGGGGTCACT CTTTCTGAT CTACATTCTG ATCTTGGGAC TTCTTTCACT  
GGCTGCAGCT CACCATGTAT GCCTGAACCA GGCTGGCCTC ATCCAAAGCC ATGCAGGATC  
ACTCAGGATT TCAGTTTCAC CCTCTATTTC CAAAGCCATT TACCTCAAAG GAQCCAGCAG  
CTACACCCCT ACAGGCTTCC AGGCACCTCA TCAGTCATGT TCCTCCTCCA TTTTACCCCC  
TACCCATCCT TGATCGGTCA TGCCTAGCCT GACCCTTTAG TAAAGCAATG AGGTAGGAAG  
AACAAACCCT TGTCCCTTTG CCATGTGGAG GAAAGTGCCT GCCTCTGGTC CGAGCCGCCT  
CGGTTCTGAA GCGAGTGCTC CTGCTTACCT TGCTCTAGGC TGTCTGCAGA AGCACCTGCC  
GGTGGCACTC AGCACCTCCT TGTGCTAGAG CCCTCCATCA CCTTCACGCT GTCCCACCAT  
GGGCCAGGAA CCAAACCAGC ACTGGGTTCT ACTGCTGTGG GGTAAACTAA CTCAGTGGAA  
TAGGGCTGGT TACTTTGGGC TGTCCAATC ATAAGTTTGG CTGCATTTTG AAAAAAGCTG  
ATCTAAATAA AGGCATGTGT ATGGCTGGTC CCCTTGTGTT TTGTTGTCTC ACATTTAGAT  
ATCAGCCATG CATGACTGAA TGGCTTCCAA TCATATACTC ACCTATCACC TACAAGAGAA  
CAATGAAAAA CACACACAAA AACAAAATCT TGAATTTTGT AATCATGCCT ATTGCTATTT  
CTTGAGCATA AGAATGGCTC AGATACTTTC CAAGACATAA AAGGAAGGCA GAGGAATAGT  
TGTTGCTGTA AAAGACATCA AGAATAAATG GGGTCATGTA CAACGGGAGG GGCCGGTTAC  
CTGAATAATG GAGTGGAGAT TGAGCTATCC TAGCTCCTCT GCTCACTAAC TGACCTGTCTG  
CATGACCGTG GACAAAACCC TGACGCGAGC TGTTTGTTTG CTAAACTTCT CTGGACCATG  
GCCTGCGGCA TATCTATAGG CATCCTGTGT TTTCCACCCA GTTTCCTTCT TCCTCGCTAA  
GCCAACGTGG AAAGGGCTGG CCGTGAATAT GCAGACAAGG TAACGAAAGT AAACCGTCAA  
TTAGTAAAAG TACTTCATTT TCCTCTTGTA TTTGCTTCAT TCTTGCTTCA CAAAGTTACG  
AAGTCCACAG CTTTATACCA AAATGTAAGA AGGCTATTTG CTTATAAACA TTTTGAGTCA  
GGTGTCATCT GATTTTATTC TTCTAATCCA TATTCAATAT TAAAAAATCA GAAACCAAGG  
GTGCTGGAGC AGCTCTAGGG CATATATTTT TCTTAAATAG GAGAAAGATT TTCAACAGCT  
TTTCCTCCTT GACCCCTCC TTTCCAATT TATTTGGGTC ACTACCTTGA ATTTAGAGTG  
AATCTGGGAA ATGTAGTCAC CAGG

(2) INFORMATION FOR SEQ ID NO: 5:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 5149 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: genomic DNA

(ix) FEATURE:  
(A) NAME/KEY: misc\_feature  
(D) OTHER INFORMATION:/label= FIGURE 2

(ix) FEATURE:  
(A) NAME/KEY: misc\_feature  
(B) POSITION : 1303..3764  
(D) OTHER INFORMATION:/note=CDS

Sub  
C1  
(ix) FEATURE:  
(A) NAME/KEY: misc\_feature  
(B) POSITION : 1631  
(D) OTHER INFORMATION:/note= CGA->TGA ; nCL1 mutation in one  
LGMD2A family

(ix) FEATURE:  
(A) NAME/KEY: misc\_feature  
(B) POSITION : 1848  
(D) OTHER INFORMATION:/note= CTG->CAG ; nCL1 mutation in one  
LGMD2A family

(ix) FEATURE:  
(A) NAME/KEY: misc\_feature  
(B) POSITION : 1853  
(D) OTHER INFORMATION:/note= CAA->CA ; nCL1 mutation in one  
LGMD2A family

(ix) FEATURE:  
(A) NAME/KEY: misc\_feature  
(B) POSITION : 2004  
(D) OTHER INFORMATION:/note= GGG->GAG ; nCL1 mutation in one  
LGMD2A family

(ix) FEATURE:  
(A) NAME/KEY: misc\_feature  
(B) POSITION : 2248  
(D) OTHER INFORMATION:/note= CGG->CG ; nCL1 mutation in one  
LGMD2A family

(ix) FEATURE:  
(A) NAME/KEY: misc\_feature  
(B) POSITION : 2364  
(D) OTHER INFORMATION:/note= GTG->GGG ; nCL1 mutation in one  
LGMD2A family

(ix) FEATURE:  
(A) NAME/KEY: misc\_feature  
(B) POSITION : 2382  
(D) OTHER INFORMATION:/note= TGG->TAG ; nCL1 mutation in one  
LGMD2A family

(ix) FEATURE:  
(A) NAME/KEY: misc\_feature  
(B) POSITION : 2771  
(D) OTHER INFORMATION:/note= CGG->TGG ; nCL1 mutation in one  
LGMD2A family

(ix) FEATURE:  
(A) NAME/KEY: misc\_feature  
(B) POSITION : 3018

(D) OTHER INFORMATION:/note= GGG->CAG ; nCL1 mutation in one  
LGMD2A family

(ix) FEATURE:

(A) NAME/KEY: misc feature

(B) POSITION : 3372..3373

(D) OTHER INFORMATION:/note= Deletion AC ; nCL1 mutation in

one

LGMD2A family

(ix) FEATURE:

(A) NAME/KEY: misc feature

(B) POSITION : 3533

(D) OTHER INFORMATION:/note= AGC->GGC ; nCL1 mutation in one

LGMD2A family

(ix) FEATURE:

(A) NAME/KEY: misc feature

(B) POSITION : 3609

(D) OTHER INFORMATION:/note= CGG->CAG ; nCL1 mutation in one

LGMD2A family

(ix) FEATURE:

(A) NAME/KEY: misc feature

(B) POSITION : 3616..3619

(D) OTHER INFORMATION:/note= Deletion AGAC ; nCL1 mutation in

one

LGMD2A family

(ix) FEATURE:

(A) NAME/KEY: misc feature

(B) POSITION : 3665..3666

(D) OTHER INFORMATION:/note= AG->TCATCT ; nCL1 mutation in one

LGMD2A family

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 5:

ATATCAGTTA GCCTGGTTTC ACTATACAGT ACATCATTTT GCTTAAAGTC ACAGCTTACG  
AGAACCTATC GATGATGTTA AGTGAGGATT TTCTCTGCTC AGGTGCACTT TTTTTTTTTT  
TTAAGACGGA GTCTCTTTCT GTCACCTGGG CTGGAGTGCA GTGGCGTGAT CTGGGTTTAC  
TACAACCTCT CCCTCCTGGG TTCAAGCAAT TCTTCTGTCT CAGCCTCCCA AGTAGCTGGG  
ATTACAGGCA CCCCCCGCCA CACCCGGGCTT ATTTTGTAT TTTTAGTAGA GACAGGGTTT  
CACTATTGTT GTCCATGCTG GTCTCGAACT CGTGACCTCA TGTGATCCAC CCGCCTCGGC  
CTCCCAAAGT GCAGAGATTA GAGACGTGAT CCACATGGCC CAGCAGGACC ACTTTTTAGC  
AGATTGAGTC CCAGTGTTCA TTTTGTGGAT GGGGAGAGAC AAGAGGTGGC AAGGTCAAGT  
GTGCAGGTAG AGACAGGGAT TTTCTCAAT GAGGACTCTG CTGAGTAGCA TTTTCCATGC  
AGACATTTCC AATGAGCGCT GACCCAAGAA CATTCTAAAA AAGATACCAA ATCTAACATT  
GAATAATGTT CTGATATCCT AAAATTTTAG GACTAAAAAT CATGTTCTCT AAAATTCACA

GAATATTTTT GTAGAATTCA GTACCTCCCG TTCACCCTAA CTAGCTTTTT TGCAATATTG  
TTTTCCATT CATTGATGGC CAGTAGTTGG GTGGTCTGTA TAACTGCCTA CTCATAAACA  
TGTCAGCAGT TCTCAGCTTC TTTCCAGTGT TCACCTTACT CAGATACTCC CTTTTCATTT  
TCTGGCAACA CCAGCACTTC ATGGCAACAG AAATGTCCCT AGCCAGGTTG TCTCTCTACC  
ATGCAGTCTC TCTTGCTCTC ATACTCACAG TGTTTCTTCA CATCTATTTT TAGTTTTCTT  
GGCTCAAGCA TCTTCAGGCC ACTGAAACAC AACCCTCACT CTCTTTCTCT CTCCCTCTGG  
CATGCATGCT GCTGGTAGGA GACCCCAAG TCAACATTGC TTCAGAAATC CTTTAGCACT  
CATTTCTCAG GAGAACTTAT GGCTTCAGAA TCACAGCTCG GTTTTTAAGA TGGACATAAC  
CTGTCCGACC TTCTGATGGG CTTTCAACTT TGAAGTGGAT GTGGACACTT TTCTCTCAGA  
TGACAGAATT ACTCCAACTT CCCCTTTGCA GTTGCTTCCT TTCCTTGAAG GTAGCTGTAT  
CTTATTTTCT TTA AAAAGCT TTTTCTTCCA AAGCCACTTG CCATGCCGAC CGTCATTAGC  
GCATCTGTGG CTCCAAGGAC AGCGGCTGAG CCCCAGTCCC CAGGGCCAGT TCCTACCCG  
GCCCAGAGCA AGGCCACTGA GGCTGGGGGT GGAAACCCAA GTGGCATCTA TTCAGCCATC  
ATCAGCCGCA ATTTTCTAT TATCGGAGTG AAAGAGAAGA CATTGAGCA ACTTCACAAG  
AAATGTCTAG AAAAGAAAGT TCTTTATGTG GACCCTGAGT TCCCACCGGA TGAGACCTCT  
CTCTTTTATA GCCAGAAGTT CCCCATCCAG TTCGTCTGGA AGAGACCTCC GGAAATTGCG  
GAGAATCCCC GATTTATCAT TGATGGAGCC AACAGAACTG ACATCTGTCA AGGAGAGCTA  
GGGGACTGCT GGTTCCTCGC AGCCATTGCC TGCCTGACCC TGAACCAGCA CCTTCTTTTC  
CGAGTCATAC CCCATGATCA AAGTTTCATC GAAAACTACG CAGGGATCTT CCACTTCCAG  
TTCTGGCGCT ATGGAGAGTG GGTGGACGTG GTTATAGATG ACTGCCTGCC AACGTACAAC  
AATCAACTGG TTTTCACCAA GTCCAACCAC CGCAATGAGT TCTGGAGTGC TCTGCTGGAG  
AAGGCTTATG CTAAGCTCCA TGGTTCCTAC GAAGCTCTGA AAGGTGGGAA CACCACAGAG  
GCCATGGAGG ACTTCACAGG AGGGGTGGCA GAGTTTTTTG AGATCAGGGA TGCTCCTAGT  
GACATGTACA AGATCATGAA GAAAGCCATC GAGAGAGGCT CCCTCATGGG CTGCTCCATT  
GATGATGGCA CGAACATGAC CTATGGAACC TCTCCTCTG GTCTGAACAT GGGGGAGTTG  
ATTGCACGGA TGGAAGGAA TATGGATAAC TCACTGCTCC AGGACTCAGA CCTCGACCCC  
AGAGGCTCAG ATGAAAGACC GACCCGGACA ATCATTCGGG TTCAGTATGA GACAAGAATG  
GCCTGCGGGC TGGTCAGAGG TCACGCCTAC TCTGTCACGG GGCTGGATGA GGTCCCGTTC  
AAAGGTGAGA AAGTGAAGCT GGTGCGGCTG CGGAATCCGT GGGGCCAGGT GGAGTGGAAC  
GGTTCTTGGA GTGATAGATG GAAGGACTGG AGCTTTGTGG ACAAAGATGA GAAGGCCCGT

Sub  
C1

CTGCAGCACC AGGTCACTGA GGATGGAGAG TTCTGGATGT CCTATGAGGA TTTCATCTAC  
CATTTACAA AGTTGGAGAT CTGCAACCTC ACGGCCGATG CTCTGCAGTC TGACAGCTT  
CAGACCTGGA CAGTGTCTGT GAACGAGGGC CGCTGGGTAC GGGGTTGCTC TGCCGGAGGC  
TGCCGCAACT TCCCAGATAC TTTCTGGACC AACCTCAGT ACCGTCTGAA GCTCCTGGAG  
GAGGACGATG ACCCTGATGA CTCGGAGGTG ATTTGCAGCT TCCTGGTGCC CCTGATGCAG  
AAGAACCGGC GGAAGGACCG GAAGCTAGGG GCCAGTCTCT TCACCATTTGG CTTGCCCATC  
TACGAGGTTC CCAAGAGAT GCACGGGAAC AAGCAGCACC TGCAGAAGGA CTTCTTCTG  
TACAACGCCT CCAAGGCCAG GAGCAAAACC TACATCAACA TCGGGGAGGT GTCCCAGCGC  
TTCCGCCTGC CTCCCAGCGA GTACGTCATC GTGCCCTCCA CCTACGAGCC CCACCAGGAG  
GGGAATTCA TCCTCCGGGT CTTCTCTGAA AAGAGGAACC TCTCTGAGGA AGTTGAAAAT  
ACCATCTCCG TGGATCGGCC AGTGAAAAAG AAAAAACCA AGCCCATCAT CTTCGTTTCG  
GACAGAGCAA ACAGCAACAA GGAGCTGGGT GTGACCAGG AGTCAGAGGA GGGCAAAGGC  
AAAACAAGCC CTGATAAGCA AAAGCAGTCC CCACAGCCAC AGCCTGGCAG CTCTGATCAG  
GAAAGTGAGG AACAGCAACA ATTCCGGAAC ATTTTCAAGC AGATAGCAGG AGATGACATG  
GAGATCTGTG CAGATGAGCT CAAGAAGTTC CTTAACACAG TCGTGAACA ACACAAGGAC  
CTGAAGACAC ACGGGTTCAC ACTGAGTCC TGCCGTAGCA TGATTGCGCT CATGGATACA  
GATGGCTCTG GAAAGCTCAA CCTGCAGGAG TTCCACCACC TCTGGAACAA GATTAAGGCC  
TGGCAGAAAA TTTTCAAACA CTATGACACA GACCAGTCCG GCACCATCAA CAGCTACGAG  
ATGCGAAATG CAGTCAACGA CGCAGGATTC CACCTCAACA ACCAGCTCTA TGACATCATT  
ACCATGCGGT ACGCAGACAA ACACATGAAC ATCGACTTTG ACAGTTTCAT CTGCTGCTTC  
GTTAGGCTGG AGGGCATGTT CAGAGCTTTT CATGCATTTG ACAAGGATGG AGATGGTATC  
ATCAAGCTCA ACCTTCTGGA GTGGCTGCAG CTCACCATGT ATGCCTGAAC CAGGCTGGCC  
TCATCCAAAG CCATGCAGGA TCACTCAGGA TTTCACTTTT ACCCTCTATT TCCAAAGCCA  
TTTACCTCAA AGGACCCAGC AGCTACACCC CTACAGGCTT CCAGGCACCT CATCAGTCAT  
GTTCTCTCTC CATTTTACCC CCTACCCATC CTTGATCGGT CATGCCTAGC CTGACCCTTT  
AGTAAACCAA TGAGGTAGGA AGAACAACCC CTTGTCCCTT TGCCATGTGG AGGAAAGTGC  
CTGCTCTGG TCCGAGCCGC CTCGGTTCTG AAGCGAGTGC TCCTGCTTAC CTTGCTCTAG  
GCTGCTGCA GAAGCACCTG CCGGTGGCAC TCAGCACCTC CTTGTGCTAG AGCCCTCCAT  
CACCTTCACG CTGTCCACC ATGGGCCAGG AACCAAACCA GCACTGGGTT CTA CTGCTGT  
GGGTAAACT AACTCAGTGG AATAGGGCTG GTTACTTTGG GCTGTCCAAC TCATAAGTTT  
GGCTGCATTT TGAAAAAGC TGATCTAAAT AAAGGCATGT GTATGGCTGG TCCCCTGTG

TTTTGTTGTC TCACATTTAG ATATCAGCCA TGCATGACTG AATGGCTTCC AATCATATAC  
TCACCTATCA CCTACAAGAG AACAAAGAAA AACACACACA AAAACAAAAT CTTGATTTT  
GTAATCATGC CTATTGCTAT TTCTTGAGCA TAAGAATGGC TCAGATACTT TCCAAGACAT  
AAAAGGAAGG CAGAGGAATA GTTGTGCTG TAAAAGACAT CAAGAATAAA TGGGTCATGT  
ACAACGGGAG GGGCCGGTTA CCTGAATAAT GGAGTGGAGA TTGAGCTATC CTAGCTCCTC  
TGCTCACTAA CTGACCTGTC GCATGACCGT GGACAAAACC CTGAACGCAG CTGTTTGT  
GCTAACTTC TCTGGACCAT GGCCTGCGGC ATATCTATAG GCATCCTGTG TTTTCCACCC  
AGTTTCCTTC TTCCTCGCTA AGCCAACGTG GAAAGGGCTG GCCGTGAATA TGCAGACAAG  
GTAACGAAAG TAAACCGTCA ATTAGTAAAA GTACTTCATT TTCCTCTTGT ATTTGCTTCA  
TATCTTGCTT CACAAAGTTA CGAAGTTCAC AGCTTTATAC CAAAATGTAA GAAGGCTATT  
TGCTTATAAA CATTTTTGCA GTCAGGTGTC ATCTGATTTT ATTCTTCTAA TCCATATTCA  
ATATTANAAA AATCAGAAAC CAAGGGTGCT GGAGCAGCTC TAGGGCATAT ATTTCTCTTA  
AATAGGAGAA AGATTTTCAA CAGCTTTTCC TCCTTGACCC CCTCCTTTCC CAATTTATTT  
GGGTCACTAC CTTGAATTTA GAGTGAATCT GGGAAATGTA GTCACCAGG

(2) INFORMATION FOR SEQ ID NO: 6:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 821 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(ix) FEATURE:

- (A) NAME/KEY: misc-feature
- (D) OTHER INFORMATION:/label= "FIGURE 2"

(ix) FEATURE:

- (A) NAME/KEY: Region
- (B) LOCATION:1..61
- (D) OTHER INFORMATION:/label= ALU  
/note= "FIGURE 3 - HUMAN"

(ix) FEATURE:

- (A) NAME/KEY: Region
- (B) LOCATION:268..309
- (D) OTHER INFORMATION:/label= ALU  
/note= "FIGURE 3 - HUMAN"

(ix) FEATURE:

- (A) NAME/KEY: Region
- (B) LOCATION:310..328
- (D) OTHER INFORMATION:/label= ALU  
/note= "FIGURE 3 - HUMAN"



(ix) FEATURE:

(A) NAME/KEY: Region  
(B) LOCATION: 578..618  
(D) OTHER INFORMATION: /label= ALU  
/note= "FIGURE 3 - HUMAN"

(ix) FEATURE:

(A) NAME/KEY: Region  
(B) LOCATION: 619..653  
(D) OTHER INFORMATION: /label= ALU  
/note= "FIGURE 3 - HUMAN"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 6:

Met Pro Thr Val Ile Ser Ala Ser Val Ala Pro Arg Thr Ala Ala Glu  
1 5 10 15  
Pro Arg Ser Pro Gly Pro Val Pro His Pro Ala Gln Ser Lys Ala Thr  
20 25 30  
Glu Ala Gly Gly Gly Asn Pro Ser Gly Ile Tyr Ser Ala Ile Ile Ser  
35 40 45  
Arg Asn Phe Pro Ile Ile Gly Val Lys Glu Lys Thr Phe Glu Gln Leu  
50 55 60  
His Lys Lys Cys Leu Glu Lys Lys Val Leu Tyr Val Asp Pro Glu Phe  
65 70 75 80  
Pro Pro Asp Glu Thr Ser Leu Phe Tyr Ser Gln Lys Phe Pro Ile Gln  
85 90 95  
Phe Val Trp Lys Arg Pro Pro Glu Ile Cys Glu Asn Pro Arg Phe Ile  
100 105 110  
Ile Asp Gly Ala Asn Arg Thr Asp Ile Cys Gln Gly Glu Leu Gly Asp  
115 120 125  
Cys Trp Phe Leu Ala Ala Ile Ala Cys Leu Thr Leu Asn Gln His Leu  
130 135 140  
Leu Phe Arg Val Ile Pro His Asp Gln Ser Phe Ile Glu Asn Tyr Ala  
145 150 155 160  
Gly Ile Phe His Phe Gln Phe Trp Arg Tyr Gly Glu Trp Val Asp Val  
165 170 175  
Val Ile Asp Asp Cys Leu Pro Thr Tyr Asn Asn Gln Leu Val Phe Thr  
180 185 190  
Lys Ser Asn His Arg Asn Glu Phe Trp Ser Ala Leu Leu Glu Lys Ala  
195 200 205  
Tyr Ala Lys Leu His Gly Ser Tyr Glu Ala Leu Lys Gly Gly Asn Thr  
210 215 220  
Thr Glu Ala Met Glu Asp Phe Thr Gly Gly Val Ala Glu Phe Phe Glu  
225 230 235 240

Ile Arg Asp Ala Pro Ser Asp Met Tyr Lys Ile Met Lys Lys Ala Ile  
 245 250 255  
 Glu Arg Gly Ser Leu Met Gly Cys Ser Ile Asp Asp Gly Thr Asn Met  
 260 265 270  
 Thr Tyr Gly Thr Ser Pro Ser Gly Leu Asn Met Gly Glu Leu Ile Ala  
 275 280 285  
 Arg Met Val Arg Asn Met Asp Asn Ser Leu Leu Gln Asp Ser Asp Leu  
 290 295 300  
 Asp Pro Arg Gly Ser Asp Glu Arg Pro Thr Arg Thr Ile Ile Pro Val  
 305 310 315 320  
 Gln Tyr Glu Thr Arg Met Ala Cys Gly Leu Val Arg Gly His Ala Tyr  
 325 330 335  
 Ser Val Thr Gly Leu Asp Glu Val Pro Phe Lys Gly Glu Lys Val Lys  
 340 345 350  
 Leu Val Arg Leu Arg Asn Pro Trp Gly Gln Val Glu Trp Asn Gly Ser  
 355 360 365  
 Trp Ser Asp Arg Trp Lys Asp Trp Ser Phe Val Asp Lys Asp Glu Lys  
 370 375 380  
 Ala Arg Leu Gln His Gln Val Thr Glu Asp Gly Glu Phe Trp Met Ser  
 385 390 395 400  
 Tyr Glu Asp Phe Ile Tyr His Phe Thr Lys Leu Glu Ile Cys Asn Leu  
 405 410 415  
 Thr Ala Asp Ala Leu Gln Ser Asp Lys Leu Gln Thr Trp Thr Val Ser  
 420 425 430  
 Val Asn Glu Gly Arg Trp Val Arg Gly Cys Ser Ala Gly Gly Cys Arg  
 435 440 445  
 Asn Phe Pro Asp Thr Phe Trp Thr Asn Pro Gln Tyr Arg Leu Lys Leu  
 450 455 460  
 Leu Glu Glu Asp Asp Asp Pro Asp Asp Ser Glu Val Ile Cys Ser Phe  
 465 470 475 480  
 Leu Val Ala Leu Met Gln Lys Asn Arg Arg Lys Asp Arg Lys Leu Gly  
 485 490 495  
 Ala Ser Leu Phe Thr Ile Gly Phe Ala Ile Tyr Glu Val Pro Lys Glu  
 500 505 510  
 Met His Gly Asn Lys Gln His Leu Gln Lys Asp Phe Phe Leu Tyr Asn  
 515 520 525  
 Ala Ser Lys Ala Arg Ser Lys Thr Tyr Ile Asn Met Arg Glu Val Ser  
 530 535 540  
 Gln Arg Phe Arg Leu Pro Pro Ser Glu Tyr Val Ile Val Pro Ser Thr  
 545 550 555 560  
 Tyr Glu Pro His Gln Glu Gly Glu Phe Ile Leu Arg Val Phe Ser Glu

sub  
 C1

Sub  
C1

565										570					575				
Lys	Arg	Asn	Leu	Ser	Glu	Glu	Val	Glu	Asn	Thr	Ile	Ser	Val	Asp	Arg				
			580					585					590						
Pro	Val	Lys	Lys	Lys	Lys	Thr	Lys	Pro	Ile	Ile	Phe	Val	Ser	Asp	Arg				
		595					600					605							
Ala	Asn	Ser	Asn	Lys	Glu	Leu	Gly	Val	Asp	Gln	Glu	Ser	Glu	Glu	Gly				
	610					615					620								
Lys	Gly	Lys	Thr	Ser	Pro	Asp	Lys	Gln	Lys	Gln	Ser	Pro	Gln	Pro	Gln				
	625				630					635					640				
Pro	Gly	Ser	Ser	Asp	Gln	Glu	Ser	Glu	Glu	Gln	Gln	Gln	Phe	Arg	Asn				
				645					650					655					
Ile	Phe	Lys	Gln	Ile	Ala	Gly	Asp	Asp	Met	Glu	Ile	Cys	Ala	Asp	Glu				
			660					665					670						
Leu	Lys	Lys	Val	Leu	Asn	Thr	Val	Val	Asn	Lys	His	Lys	Asp	Leu	Lys				
		675					680					685							
Thr	His	Gly	Phe	Thr	Leu	Glu	Ser	Cys	Arg	Ser	Met	Ile	Ala	Leu	Met				
	690					695					700								
Asp	Thr	Asp	Gly	Ser	Gly	Lys	Leu	Asn	Leu	Gln	Glu	Phe	His	His	Leu				
	705				710					715					720				
Trp	Asn	Lys	Ile	Lys	Ala	Trp	Gln	Lys	Ile	Phe	Lys	His	Tyr	Asp	Thr				
			725						730					735					
Asp	Gln	Ser	Gly	Thr	Ile	Asn	Ser	Tyr	Glu	Met	Arg	Asn	Ala	Val	Asn				
			740					745					750						
Asp	Ala	Gly	Phe	His	Leu	Asn	Asn	Gln	Leu	Tyr	Asp	Ile	Ile	Thr	Met				
		755				760						765							
Arg	Tyr	Ala	Asp	Lys	His	Met	Asn	Ile	Asp	Phe	Asp	Ser	Phe	Ile	Cys				
	770					775					780								
Cys	Phe	Val	Arg	Leu	Glu	Gly	Met	Phe	Arg	Ala	Phe	His	Ala	Phe	Asp				
	785				790					795					800				
Lys	Asp	Gly	Asp	Gly	Ile	Ile	Lys	Leu	Asn	Val	Leu	Glu	Trp	Leu	Gln				
				805					810					815					
Leu	Thr	Met	Tyr	Ala															
			820																

(2) INFORMATION FOR SEQ ID NO: 7:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 821 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(ix) FEATURE:

- (A) NAME/KEY: Region
- (B) LOCATION:1..61
- (D) OTHER INFORMATION:/label= ALU

(ix) FEATURE:

- (A) NAME/KEY: Region
- (B) LOCATION:268..309
- (D) OTHER INFORMATION:/label= ALU  
/note= "FIGURE 3 - RAT"

(ix) FEATURE:

- (A) NAME/KEY: Region
- (B) LOCATION:310..328
- (D) OTHER INFORMATION:/label= ALU  
/note= "FIGURE 3 - RAT"

(ix) FEATURE:

- (A) NAME/KEY: Region
- (B) LOCATION:578..618
- (D) OTHER INFORMATION:/label= ALU  
/note= "FIGURE 3 - RAT"

(ix) FEATURE:

- (A) NAME/KEY: Region
- (B) LOCATION:619..653
- (D) OTHER INFORMATION:/label= ALU  
/note= "FIGURE 3 - RAT"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 7:

Met Pro Thr Val Ile Ser Pro Thr Val Ala Pro Arg Thr Gly Ala Glu  
1 5 10 15  
Pro Arg Ser Pro Gly Pro Val Pro His Pro Ala Gln Gly Lys Thr Thr  
20 25 30  
Glu Ala Gly Gly Gly His Pro Gly Gly Ile Tyr Ser Ala Ile Ile Ser  
35 40 45  
Arg Asn Phe Pro Ile Ile Gly Val Lys Glu Lys Thr Phe Glu Gln Leu  
50 55 60  
His Lys Lys Cys Leu Glu Lys Lys Val Leu Tyr Leu Asp Pro Glu Phe  
65 70 75 80  
Pro Pro Asp Glu Thr Ser Leu Phe Tyr Ser Gln Lys Phe Pro Ile Gln  
85 90 95  
Phe Val Trp Lys Arg Pro Pro Glu Ile Cys Glu Asn Pro Arg Phe Ile  
100 105 110  
Ile Gly Gly Ala Asn Arg Thr Asp Ile Cys Gln Gly Asp Leu Gly Asp  
115 120 125  
Cys Trp Leu Leu Ala Ala Ile Ala Cys Leu Thr Leu Asn Glu Arg Leu  
130 135 140  
Leu Phe Arg Val Ile Pro His Asp Gln Ser Phe Thr Glu Asn Tyr Ala  
145 150 155 160

Sub  
C1

Gly Ile Phe His Phe Gln Phe Trp Arg Tyr Gly Asp Trp Val Asp Val  
165 170 175  
Val Ile Asp Asp Cys Leu Pro Thr Tyr Asn Asn Gln Leu Val Phe Thr  
180 185 190  
Lys Ser Asn His Arg Asn Glu Phe Trp Ser Ala Leu Leu Glu Lys Ala  
195 200 205  
Tyr Ala Lys Leu His Gly Ser Tyr Glu Ala Leu Lys Gly Gly Asn Thr  
210 215 220  
Thr Glu Ala Met Glu Asp Phe Thr Gly Gly Val Thr Glu Phe Phe Glu  
225 230 235 240  
Ile Lys Asp Ala Pro Ser Asp Met Tyr Lys Ile Met Arg Lys Ala Ile  
245 250 255  
Glu Arg Gly Ser Leu Met Gly Cys Ser Ile Asp Asp Gly Thr Asn Met  
260 265 270  
Thr Tyr Gly Thr Ser Pro Ser Gly Leu Asn Met Gly Glu Leu Ile Ala  
275 280 285  
Arg Met Val Arg Asn Met Asp Asn Ser Leu Leu Arg Asp Ser Asp Leu  
290 295 300  
Asp Pro Arg Ala Ser Asp Asp Arg Pro Ser Arg Thr Ile Val Pro Val  
305 310 315 320  
Gln Tyr Glu Thr Arg Met Ala Cys Gly Leu Val Arg Gly His Ala Tyr  
325 330 335  
Ser Val Thr Gly Leu Glu Glu Ala Leu Phe Lys Gly Glu Lys Val Lys  
340 345 350  
Leu Val Arg Leu Arg Asn Pro Trp Gly Gln Val Glu Trp Asn Gly Ser  
355 360 365  
Trp Ser Asp Gly Trp Lys Asp Trp Ser Phe Val Asp Lys Asp Glu Lys  
370 375 380  
Ala Arg Leu Gln His Gln Val Thr Glu Asp Gly Glu Phe Trp Met Ser  
385 390 395 400  
Tyr Asp Asp Phe Val Tyr His Phe Thr Lys Leu Glu Ile Cys Asn Leu  
405 410 415  
Thr Ala Asp Ala Leu Glu Ser Asp Lys Leu Gln Thr Trp Thr Val Ser  
420 425 430  
Val Asn Glu Gly Arg Trp Val Arg Gly Cys Ser Ala Gly Gly Cys Arg  
435 440 445  
Asn Phe Pro Asp Thr Phe Trp Thr Asn Pro Gln Tyr Arg Leu Lys Leu  
450 455 460  
Leu Glu Glu Asp Asp Asp Pro Asp Asp Ser Glu Val Ile Cys Ser Phe  
465 470 475 480

Sub  
C1

Leu Val Ala Leu Met Gln Lys Asn Arg Arg Lys Asp Arg Lys Leu Gly  
485 490 495

Ala Asn Leu Phe Thr Ile Gly Phe Ala Ile Tyr Glu Val Pro Lys Glu  
500 505 510

Met His Gly Asn Lys Gln His Leu Gln Lys Asp Phe Phe Leu Tyr Asn  
515 520 525

Ala Ser Lys Ala Arg Ser Lys Thr Tyr Ile Asn Met Arg Glu Val Ser  
530 535 540

Gln Arg Phe Arg Leu Pro Pro Ser Glu Tyr Val Ile Val Pro Ser Thr  
545 550 555 560

Tyr Glu Pro His Gln Glu Gly Glu Phe Ile Leu Arg Val Phe Ser Glu  
565 570 575

Lys Arg Asn Leu Ser Glu Glu Ala Glu Asn Thr Ile Ser Val Asp Arg  
580 585 590

Pro Val Lys Lys Lys Lys Asn Lys Pro Ile Ile Phe Val Ser Asp Arg  
595 600 605

Ala Asn Ser Asn Lys Glu Leu Gly Val Asp Gln Glu Ala Glu Glu Gly  
610 615 620

Lys Asp Lys Thr Gly Pro Asp Lys Gln Gly Glu Ser Pro Gln Pro Arg  
625 630 635 640

Pro Gly His Thr Asp Gln Glu Ser Glu Glu Gln Gln Gln Phe Arg Asn  
645 650 655

Ile Phe Arg Gln Ile Ala Gly Asp Asp Met Glu Ile Cys Ala Asp Glu  
660 665 670

Leu Lys Asn Val Leu Asn Thr Val Val Asn Lys His Lys Asp Leu Lys  
675 680 685

Thr Gln Gly Phe Thr Leu Glu Ser Cys Arg Ser Met Ile Ala Leu Met  
690 695 700

Asp Thr Asp Gly Ser Gly Arg Leu Asn Leu Gln Glu Phe His His Leu  
705 710 715 720

Trp Lys Lys Ile Lys Ala Trp Gln Lys Ile Phe Lys His Tyr Asp Thr  
725 730 735

Asp His Ser Gly Thr Ile Asn Ser Tyr Glu Met Arg Asn Ala Val Asn  
740 745 750

Asp Ala Gly Phe His Leu Asn Ser Gln Leu Tyr Asp Ile Ile Thr Met  
755 760 765

Arg Tyr Ala Asp Lys His Met Asn Ile Asp Phe Asp Ser Phe Ile Cys  
770 775 780

Cys Phe Val Arg Leu Glu Gly Met Phe Arg Ala Phe His Ala Phe Asp  
785 790 795 800

Lys Asp Gly Asp Gly Ile Ile Lys Leu Asn Val Leu Glu Trp Leu Gln

805

810

815

Leu Thr Met Tyr Ala  
820

(2) INFORMATION FOR SEQ ID NO: 8:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 821 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(ix) FEATURE:

- (A) NAME/KEY: Region
- (B) LOCATION:1..61
- (D) OTHER INFORMATION:/label= ALU  
/note= "FIGURE 3 - PIG"

(ix) FEATURE:

- (A) NAME/KEY: Region
- (B) LOCATION:268..309
- (D) OTHER INFORMATION:/label= ALU  
/note= "FIGURE 3 - PIG"

(ix) FEATURE:

- (A) NAME/KEY: Region
- (B) LOCATION:310..328
- (D) OTHER INFORMATION:/label= ALU  
/note= "FIGURE 3 - PIG"

(ix) FEATURE:

- (A) NAME/KEY: Region
- (B) LOCATION:578..618
- (D) OTHER INFORMATION:/label= ALU  
/note= "FIGURE 3 - PIG"

(ix) FEATURE:

- (A) NAME/KEY: Region
- (B) LOCATION:619..653
- (D) OTHER INFORMATION:/label= ALI  
/note= "FIGURE 3 - PIG"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 8:

Met	Pro	Thr	Val	Ile	Ser	Ala	Ser	Val	Ala	Pro	Arg	Thr	Ala	Ala	Glu
1				5				10					15		
Pro	Arg	Ser	Pro	Gly	Pro	Val	Pro	His	Pro	Ala	Gln	Ser	Lys	Ala	Thr
			20				25					30			
Glu	Ala	Gly	Gly	Gly	Asn	Pro	Ser	Gly	Ile	Tyr	Ser	Ala	Ile	Ile	Ser
		35				40					45				
Arg	Asn	Phe	Pro	Ile	Ile	Gly	Val	Lys	Glu	Lys	Thr	Phe	Glu	Gln	Leu
	50					55					60				

Sub  
C1

Sub  
C1

His Lys Lys Cys Leu Glu Lys Lys Val Leu Tyr Val Asp Pro Glu Phe  
65 70 75 80  
Pro Pro Asp Glu Thr Ser Leu Phe Tyr Ser Gln Lys Phe Pro Ile Gln  
85 90 95  
Phe Val Trp Lys Arg Pro Pro Glu Ile Cys Glu Asn Pro Arg Phe Ile  
100 105 110  
Ile Asp Gly Ala Asn Arg Thr Asp Ile Cys Gln Gly Glu Leu Gly Asp  
115 120 125  
Cys Trp Phe Leu Ala Ala Ile Ala Cys Leu Thr Leu Asn Gln His Leu  
130 135 140  
Leu Phe Arg Val Ile Pro His Asp Gln Ser Phe Ile Glu Asn Tyr Ala  
145 150 155 160  
Gly Ile Phe His Phe Gln Phe Trp Arg Tyr Gly Glu Trp Val Asp Val  
165 170 175  
Val Ile Asp Asp Cys Leu Pro Thr Tyr Asn Asn Gln Leu Val Phe Thr  
180 185 190  
Lys Ser Asn His Arg Asn Glu Phe Trp Ser Ala Leu Leu Glu Lys Ala  
195 200 205  
Tyr Ala Lys Leu His Gly Ser Tyr Glu Ala Leu Lys Gly Gly Asn Thr  
210 215 220  
Thr Glu Ala Met Glu Asp Phe Thr Gly Gly Val Ala Glu Phe Phe Glu  
225 230 235 240  
Ile Arg Asp Ala Pro Ser Asp Met Tyr Lys Ile Met Lys Lys Ala Ile  
245 250 255  
Glu Arg Gly Ser Leu Met Gly Cys Ser Ile Asp Asp Gly Thr Asn Met  
260 265 270  
Thr Tyr Gly Thr Ser Pro Ser Gly Leu Asn Met Gly Asp Leu Ile Ala  
275 280 285  
Arg Met Val Arg Asn Met Glu Asn Ser Arg Leu Arg Asp Ser Ile Leu  
290 295 300  
Asp Pro Glu Val Ser Asp Asp Arg Pro Thr Arg Thr Ile Val Pro Val  
305 310 315 320  
Gln Phe Glu Thr Arg Met Ala Cys Gly Leu Val Arg Gly His Ala Tyr  
325 330 335  
Ser Val Thr Gly Leu Glu Glu Ala Leu Phe Lys Gly Glu Lys Val Lys  
340 345 350  
Leu Val Arg Leu Arg Asn Pro Trp Gly Gln Val Glu Trp Asn Gly Ser  
355 360 365  
Trp Ser Asp Ser Trp Lys Asp Trp Ser Phe Val Asp Lys Asp Glu Lys  
370 375 380  
Ala Arg Leu Gln His Gln Val Thr Glu Asp Gly Glu Phe Trp Met Ser



Sub  
C1

385	390	395	400
Tyr Asp Asp Phe Ile Tyr His Phe Thr Lys Leu Glu Ile Cys Asn Leu	405	410	415
Thr Ala Asp Ala Leu Glu Ser Asp Lys Leu Gln Thr Trp Thr Val Ser	420	425	430
Val Asn Glu Gly Arg Trp Val Arg Gly Cys Ser Ala Gly Thr Gly Arg	435	440	445
Asn Phe Pro Asp Thr Phe Trp Thr Asn Pro Gln Tyr Arg Leu Lys Leu	450	455	460
Leu Glu Glu Asp Asp Asp Pro Asp Asp Ser Glu Val Ile Cys Ser Phe	465	470	475
Leu Val Ala Leu Met Gln Arg Asn Arg Arg Lys Asp Arg Lys Leu Gly	485	490	495
Ala Asn Leu Phe Thr Ile Gly Phe Ala Ile Tyr Glu Val Pro Lys Glu	500	505	510
Met His Gly Asn Lys Gln His Leu Gln Lys Asp Phe Phe Leu Tyr Asn	515	520	525
Ala Ser Lys Ala Arg Ser Arg Thr Tyr Ile Asn Met Arg Glu Val Ser	530	535	540
Glu Arg Phe Arg Leu Pro Pro Ser Glu Tyr Val Ile Val Pro Ser Thr	545	550	555
Tyr Glu Pro His Gln Glu Gly Glu Phe Met Leu Arg Val Phe Ser Glu	565	570	575
Lys Arg Lys Leu Ser Glu Glu Val Glu Asn Thr Ile Ser Val Asp Arg	580	585	590
Pro Val Arg Lys Lys Lys Thr Lys Pro Ile Ile Phe Val Ser Asp Arg	595	600	605
Ala Asn Ser Asn Lys Glu Leu Gly Val Asp Gln Glu Ser Glu Glu Gly	610	615	620
Gln Asp Lys Thr Ser Pro Asp Lys Gln Glu Lys Ser Pro Lys Pro Glu	625	630	635
Pro Ser Asn Thr Asp Gln Glu Ser Glu Glu Gln Gln Gln Phe Arg Asn	645	650	655
Ile Phe Lys Gln Ile Ala Gly Asp Asp Met Glu Ile Cys Ala Asp Glu	660	665	670
Leu Lys Lys Val Leu Asn Thr Val Val Asn Lys His Lys Asp Leu Lys	675	680	685
Thr His Gly Phe Thr Leu Glu Ser Cys Arg Ser Met Ile Ala Leu Met	690	695	700
Asp Thr Asp Gly Ser Gly Lys Leu Asn Leu Gln Glu Phe His His Leu	705	710	715
			720

Trp Asn Lys Ile Lys Ala Trp Gln Lys Ile Phe Lys His Tyr Asp Thr  
 725 730 735  
 Asp Gln Ser Gly Thr Ile Asn Ser Tyr Glu Met Arg Asn Ala Val Asn  
 740 745 750  
 Asp Ala Gly Phe His Leu Asn Asn Gln Leu Tyr Asp Ile Ile Thr Met  
 755 760 765  
 Arg Tyr Ala Asp Lys His Met Asn Ile Asp Phe Asp Ser Phe Ile Cys  
 770 775 780  
 Cys Phe Val Arg Leu Glu Gly Met Phe Arg Ala Phe His Ala Phe Asp  
 785 790 795 800  
 Lys Asp Gly Asp Gly Ile Ile Lys Leu Asn Val Leu Glu Trp Leu Gln  
 805 810 815  
 Leu Thr Met Tyr Ala  
 820

(2) INFORMATION FOR SEQ ID NO: 9:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 821 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(ix) FEATURE:

- (A) NAME/KEY: Region
- (B) LOCATION:1..61
- (D) OTHER INFORMATION:/label= ALU  
/note= "FIGURE 3 - COW"

(ix) FEATURE:

- (A) NAME/KEY: Region
- (B) LOCATION:268..309
- (D) OTHER INFORMATION:/label= ALU  
/note= "FIGURE 3 - COW"

(ix) FEATURE:

- (A) NAME/KEY: Region
- (B) LOCATION:310..328
- (D) OTHER INFORMATION:/label= ALU  
/note= "FIGURE 3 - COW"

(ix) FEATURE:

- (A) NAME/KEY: Region
- (B) LOCATION:578..618
- (D) OTHER INFORMATION:/label= ALU  
/note= "FIGURE 3 - COW"

(ix) FEATURE:

- (A) NAME/KEY: Region
- (B) LOCATION:619..653
- (D) OTHER INFORMATION:/label= ALU

/note= "FIGURE 3 - COW"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 9:

Met Pro Thr Val Ile Ser Ala Ser Val Ala Pro Arg Thr Ala Ala Glu  
1 5 10 15  
Pro Arg Ser Pro Gly Pro Val Pro His Pro Ala Gln Ser Lys Ala Thr  
20 25 30  
Glu Ala Gly Gly Gly Asn Pro Ser Gly Ile Tyr Ser Ala Ile Ile Ser  
35 40 45  
Arg Asn Phe Pro Ile Ile Gly Val Lys Glu Lys Thr Phe Glu Gln Leu  
50 55 60  
His Lys Lys Cys Leu Glu Lys Lys Val Leu Tyr Val Asp Pro Glu Phe  
65 70 75 80  
Pro Pro Asp Glu Thr Ser Leu Phe Tyr Ser Gln Lys Phe Pro Ile Gln  
85 90 95  
Phe Val Trp Lys Arg Pro Pro Glu Ile Cys Glu Asn Pro Arg Phe Ile  
100 105 110  
Ile Asp Gly Ala Asn Arg Thr Asp Ile Cys Gln Gly Glu Leu Gly Asp  
115 120 125  
Cys Trp Phe Leu Ala Ala Ile Ala Cys Leu Thr Leu Asn Gln His Leu  
130 135 140  
Leu Phe Arg Val Ile Pro His Asp Gln Ser Phe Ile Glu Asn Tyr Ala  
145 150 155 160  
Gly Ile Phe His Phe Gln Phe Trp Arg Tyr Gly Glu Trp Val Asp Val  
165 170 175  
Val Ile Asp Asp Cys Leu Pro Thr Tyr Asn Asn Gln Leu Val Phe Thr  
180 185 190  
Lys Ser Asn His Arg Asn Glu Phe Trp Ser Ala Leu Leu Glu Lys Ala  
195 200 205  
Tyr Ala Lys Leu His Gly Ser Tyr Glu Ala Leu Lys Gly Gly Asn Thr  
210 215 220  
Thr Glu Ala Met Glu Asp Phe Thr Gly Gly Val Ala Glu Phe Phe Glu  
225 230 235 240  
Ile Arg Asp Ala Pro Ser Asp Met Tyr Lys Ile Met Lys Lys Ala Ile  
245 250 255  
Glu Arg Gly Ser Leu Met Gly Cys Ser Ile Asp Asp Gly Thr Asn Met  
260 265 270  
Thr Tyr Gly Thr Ser Pro Ser Gly Leu Asn Met Gly Glu Leu Ile Glu  
275 280 285  
Arg Met Val Arg Asn Met Asp Asn Ser Arg Leu Arg Asp Ser Ile Leu  
290 295 300

Sub  
C1

Asp Pro Glu Val Ser Asp Asp Arg Pro Thr Arg Met Ile Val Pro Val  
305 310 315 320  
Gln Phe Glu Thr Arg Met Ala Cys Gly Leu Val Arg Gly His Ala Tyr  
325 330 335  
Ser Val Thr Gly Leu Glu Glu Ala Leu Tyr Lys Gly Glu Lys Val Lys  
340 345 350  
Leu Val Arg Leu Arg Asn Pro Trp Gly Gln Val Glu Trp Asn Gly Ser  
355 360 365  
Trp Ser Asp Ser Trp Lys Asp Trp Ser Tyr Val Asp Lys Asp Glu Lys  
370 375 380  
Ala Arg Leu Gln His Gln Val Thr Glu Asp Gly Glu Phe Trp Met Ser  
385 390 395 400  
Tyr Glu Asp Phe Ile Tyr His Phe Thr Lys Leu Glu Ile Cys Asn Leu  
405 410 415  
Thr Ala Asp Ala Leu Gln Ser Asp Lys Leu Gln Thr Trp Thr Val Ser  
420 425 430  
Val Asn Glu Gly Arg Trp Val Arg Gly Cys Ser Ala Gly Gly Cys Arg  
435 440 445  
Asn Phe Pro Asp Thr Phe Trp Thr Asn Pro Gln Tyr Arg Leu Lys Leu  
450 455 460  
Leu Glu Glu Asp Asp Asp Pro Asp Asp Ser Glu Val Ile Cys Ser Phe  
465 470 475 480  
Leu Val Ala Leu Met Gln Lys Asn Arg Arg Lys Asp Arg Lys Leu Gly  
485 490 495  
Ala Ser Leu Phe Thr Ile Gly Phe Ala Ile Tyr Glu Val Pro Lys Glu  
500 505 510  
Met His Gly Asn Lys Gln His Leu Gln Lys Asp Phe Phe Leu Tyr Asn  
515 520 525  
Ala Ser Lys Ala Arg Ser Lys Thr Tyr Ile Asn Met Arg Glu Val Ser  
530 535 540  
Gln Arg Phe Arg Leu Pro Pro Ser Glu Tyr Val Ile Val Pro Ser Thr  
545 550 555 560  
Tyr Glu Pro His Gln Glu Gly Glu Phe Ile Leu Arg Val Phe Ser Glu  
565 570 575  
Lys Arg Asn Leu Ser Glu Glu Val Glu Asn Thr Ile Ser Val Asp Arg  
580 585 590  
Pro Val Lys Lys Lys Lys Thr Lys Pro Ile Ile Phe Val Ser Asp Arg  
595 600 605  
Ala Asn Ser Asn Lys Glu Leu Gly Val Asp Gln Glu Ser Glu Glu Gly  
610 615 620

Lys Gly Lys Thr Ser Pro Asp Lys Gln Lys Gln Ser Pro Gln Pro Gln  
 625 630 635 640  
 Pro Gly Ser Ser Asp Gln Glu Ser Glu Glu Gln Gln Gln Phe Arg Asn  
 645 650 655  
 Ile Phe Lys Gln Ile Ala Gly Asp Asp Met Glu Ile Cys Ala Asp Glu  
 660 665 670  
 Leu Lys Lys Val Leu Asn Thr Val Val Asn Lys His Lys Asp Leu Lys  
 675 680 685  
 Thr His Gly Phe Thr Leu Glu Ser Cys Arg Ser Met Ile Ala Leu Met  
 690 695 700  
 Asp Thr Asp Gly Ser Gly Lys Leu Asn Leu Gln Glu Phe His His Leu  
 705 710 715 720  
 Trp Asn Lys Ile Lys Ala Trp Gln Lys Ile Phe Lys His Tyr Asp Thr  
 725 730 735  
 Asp Gln Ser Gly Thr Ile Asn Ser Tyr Glu Met Arg Asn Ala Val Asn  
 740 745 750  
 Asp Ala Gly Phe His Leu Asn Asn Gln Leu Tyr Asp Ile Ile Thr Met  
 755 760 765  
 Arg Tyr Ala Asp Lys His Met Asn Ile Asp Phe Asp Ser Phe Ile Cys  
 770 775 780  
 Cys Phe Val Arg Leu Glu Gly Met Phe Arg Ala Phe His Ala Phe Asp  
 785 790 795 800  
 Lys Asp Gly Asp Gly Ile Ile Lys Leu Asn Val Leu Glu Trp Leu Gln  
 805 810 815  
 Leu Thr Met Tyr Ala  
 820

(2) INFORMATION FOR SEQ ID NO: 10:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 20 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: double
  - (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: cDNA
- (ix) FEATURE:
  - (A) NAME/KEY: misc\_feature
  - (D) OTHER INFORMATION: /label= Table 1
- (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 10:

ATGGAGCCAA CAGAACTGAC

(2) INFORMATION FOR SEQ ID NO: 11:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 21 base pairs

(B) TYPE: nucleic acid  
(C) STRANDEDNESS: double  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Primer

(ix) FEATURE:

(A) NAME/KEY: misc\_feature  
(D) OTHER INFORMATION: /label= Table 1

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 11:

GTATGACTCG GAAAAGAAGGT

(2) INFORMATION FOR SEQ ID NO: 12:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 20 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: double  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Primer

(ix) FEATURE:

(A) NAME/KEY: misc\_feature  
(D) OTHER INFORMATION: /label= Table 1

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 12:

TAAGCAAAAG CAGTCCCCAC

(2) INFORMATION FOR SEQ ID NO: 13:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 21 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: double  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: primer

(ix) FEATURE:

(A) NAME/KEY: misc\_feature  
(D) OTHER INFORMATION: /label= Table 1

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 13:

TTGCTGTTCC TCACTTTCCTG

(2) INFORMATION FOR SEQ ID NO: 14:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 20 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: double  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: primer

(ix) FEATURE:

- (A) NAME/KEY: misc\_feature  
(D) OTHER INFORMATION:/label= Table 1

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 14:

GTTTCATCTG CTGCTTCGTT

(2) INFORMATION FOR SEQ ID NO: 15:

- (i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 20 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: double  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: primer

(ix) FEATURE:

- (A) NAME/KEY: misc\_feature  
(D) OTHER INFORMATION:/label= Table 1

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 15:

CTGGTTCAGG CATACTGGT

(2) INFORMATION FOR SEQ ID NO: 16:

- (i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 22 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: double  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: primer

(ix) FEATURE:

- (A) NAME/KEY: misc\_feature  
(D) OTHER INFORMATION:/label= Table 1

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 16:

TTCTTTATGT GGACCCTGAG TT

(2) INFORMATION FOR SEQ ID NO: 17:

- (i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 19 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: double  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: primer

(ix) FEATURE:

- (A) NAME/KEY: misc\_feature  
(D) OTHER INFORMATION:/label= Table 1

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 17:

ACGAACTGGA TGGGGAAC

(2) INFORMATION FOR SEQ ID NO:18 :

- (i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 20 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: double  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: primer

- (ix) FEATURE:  
(A) NAME/KEY: misc feature  
(D) OTHER INFORMATION:/label= Table 3

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:18:

TTCAGTACCTCCCGTTCACC

(2) INFORMATION FOR SEQ ID NO:19:

- (i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 20 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: double  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: primer

- (ix) FEATURE:  
(A) NAME/KEY: misc feature  
(D) OTHER INFORMATION:/label= Table 3

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:19:

GATGCTTGAGCCAGGAAAAC

(2) INFORMATION FOR SEQ ID NO:20:

- (i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 22 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: double  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: primer

- (ix) FEATURE:  
(A) NAME/KEY: misc feature  
(D) OTHER INFORMATION:/label= Table 3

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:20:

CTTTCCTTGAAGGTAGCTGTAT

(2) INFORMATION FOR SEQ ID NO:21 :

- (i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 20 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: double



(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: primer

(ix) FEATURE:

(A) NAME/KEY: misc\_feature

(D) OTHER INFORMATION:/label= Table 3

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:21 :

GAGGTGCTGAGTGAGAGGAC

(2) INFORMATION FOR SEQ ID NO:22 :

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 22 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: double

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: primer

(ix) FEATURE:

(A) NAME/KEY: misc\_feature

(D) OTHER INFORMATION:/label= Table 3

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:22 :

ACTCCGTCTCAAAAAAATACCT

(2) INFORMATION FOR SEQ ID NO:23:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 20 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: double

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: primer

(ix) FEATURE:

(A) NAME/KEY: misc\_feature

(D) OTHER INFORMATION:/label= Table 3

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:23 :

ATTGTCCCTTTAOCCTCCTGG

(2) INFORMATION FOR SEQ ID NO:24 :

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 20 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: double

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: primer

(ix) FEATURE:

(A) NAME/KEY: misc\_feature

(D) OTHER INFORMATION:/label= Table 3

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:24 :

TGGAAGTAGGAGAGTGGGCA

(2) INFORMATION FOR SEQ ID NO:25 :

(i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 20 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: double  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: primer

(ix) FEATURE:  
(A) NAME/KEY: misc feature  
(D) OTHER INFORMATION:/label= Table 3

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:25 :

GGGTAGATGGGTGGGAAGTT

(2) INFORMATION FOR SEQ ID NO:26 :

(i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 20 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: double  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: primer

(ix) FEATURE:  
(A) NAME/KEY: misc feature  
(D) OTHER INFORMATION:/label= Table 3

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:26 :

GAGGAATGTGGAGGAAGGAC

(2) INFORMATION FOR SEQ ID NO:27:

(i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 20 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: double  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: primer

(ix) FEATURE:  
(A) NAME/KEY: misc feature  
(D) OTHER INFORMATION:/label= Table 3

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:27:

TTCCTGTGAGTGAGGTCTCG

(2) INFORMATION FOR SEQ ID NO:28 :

Sub  
C1

- (i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 20 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: double  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: primer

- (ix) FEATURE:  
(A) NAME/KEY: misc\_feature  
(D) OTHER INFORMATION:/label= Table 3

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:28 :

GGAAGCTCTGTGACCCCAAAT

(2) INFORMATION FOR SEQ ID NO:29 :

- (i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 20 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: double  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: primer

- (ix) FEATURE:  
(A) NAME/KEY: misc\_feature  
(D) OTHER INFORMATION:/label= Table 3

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:29 :

TCCTCAAACAAAACATTTCGC

(2) INFORMATION FOR SEQ ID NO:30 :

- (i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 20 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: double  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: primer

- (ix) FEATURE:  
(A) NAME/KEY: misc\_feature  
(D) OTHER INFORMATION:/label= Table 3

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:30:

GTTCCCTACATTCTCCATCG

(2) INFORMATION FOR SEQ ID NO:31 :

- (i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 21 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: double  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: primer

(ix) FEATURE:

- (A) NAME/KEY: misc\_feature
- (D) OTHER INFORMATION:/label= Table 3

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:31 :

GTTATTTCAACCCAGACCCTT

(2) INFORMATION FOR SEQ ID NO:32 :

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 21 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: primer

(ix) FEATURE:

- (A) NAME/KEY: misc\_feature
- (D) OTHER INFORMATION:/label= Table 3

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:32 :

AATGGGTTCTCTGGTTACTGC

(2) INFORMATION FOR SEQ ID NO:33 :

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 21 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: primer

(ix) FEATURE:

- (A) NAME/KEY: misc\_feature
- (D) OTHER INFORMATION:/label= Table 3

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:33 :

AGCACGAAAAGCAAAGATAAA

(2) INFORMATION FOR SEQ ID NO:34 :

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 20 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: primer

(ix) FEATURE:

- (A) NAME/KEY: misc\_feature
- (D) OTHER INFORMATION:/label= Table 3

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:34 :

GTAAGAGATTTGCCCCCAG

(2) INFORMATION FOR SEQ ID NO:35 :

- (i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 20 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: double  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: primer

- (ix) FEATURE:  
(A) NAME/KEY: misc\_feature  
(D) OTHER INFORMATION:/label= Table 3

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:35 :

TCTGCGGATCATTGGTTTTG

(2) INFORMATION FOR SEQ ID NO:36 :

- (i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 20 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: double  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: primer

- (ix) FEATURE:  
(A) NAME/KEY: misc\_feature  
(D) OTHER INFORMATION:/label= Table 3

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:36 :

CCTTCCCTTCTTCTGCTTC

(2) INFORMATION FOR SEQ ID NO:37 :

- (i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 20 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: double  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: primer

- (ix) FEATURE:  
(A) NAME/KEY: misc\_feature  
(D) OTHER INFORMATION:/label= Table 3

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:37 :

CTCTCTTCCCCACCCTTACC

(2) INFORMATION FOR SEQ ID NO:38 :

- (i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 20 base pairs  
(B) TYPE: nucleic acid

Sub  
C1

(C) STRANDEDNESS: double  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: primer

(ix) FEATURE:

(A) NAME/KEY: misc\_feature  
(D) OTHER INFORMATION:/label= Table 3

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:38 :

CCTCCTCACCTGCTCCATA

(2) INFORMATION FOR SEQ ID NO:39 :

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 20 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: double  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: primer

(ix) FEATURE:

(A) NAME/KEY: misc\_feature  
(D) OTHER INFORMATION:/label= Table 3

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:39 :

TTTTTCGGCTTAGACCCTCC

(2) INFORMATION FOR SEQ ID NO:40 :

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 22 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: double  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: primer

(ix) FEATURE:

(A) NAME/KEY: misc\_feature  
(D) OTHER INFORMATION:/label= Table 3

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:40 :

TGTGGGGAATAGAAATAATGG

(2) INFORMATION FOR SEQ ID NO:41 :

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 19 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: double  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: primer

(ix) FEATURE:

(A) NAME/KEY: misc\_feature

(D) OTHER INFORMATION:/label= Table 3

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:41 :

CCAGGAGCTCTGTGGGTCA

(2) INFORMATION FOR SEQ ID NO:42 :

- (i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 21 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: double  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: primer

- (ix) FEATURE:  
(A) NAME/KEY: misc\_feature  
(D) OTHER INFORMATION:/label= Table 3

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:42 :

GGCTCCTCATCCTCATTACA

(2) INFORMATION FOR SEQ ID NO:43 :

- (i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 20 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: double  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: primer

- (ix) FEATURE:  
(A) NAME/KEY: misc\_feature  
(D) OTHER INFORMATION:/label= Table 3

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:43:

GTGGAGGAGGGTGAGTGTGC

(2) INFORMATION FOR SEQ ID NO:44 :

- (i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 20 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: double  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: primer

- (ix) FEATURE:  
(A) NAME/KEY: misc\_feature  
(D) OTHER INFORMATION:/label= Table 3

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:44 :

TGTGGCAGGACAGGACGTTC

(2) INFORMATION FOR SEQ ID NO:45 :

Sub  
C1

- Sub  
C1
- (i) SEQUENCE CHARACTERISTICS:
    - (A) LENGTH: 20 base pairs
    - (B) TYPE: nucleic acid
    - (C) STRANDEDNESS: double
    - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: primer

(ix) FEATURE:

- (A) NAME/KEY: misc\_feature
- (D) OTHER INFORMATION:/label= Table 3

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:45 :

TTCAACCTCTGGAGTGGGCC

(2) INFORMATION FOR SEQ ID NO:46 :

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 20 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: double
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: primer

(ix) FEATURE:

- (A) NAME/KEY: misc\_feature
- (D) OTHER INFORMATION:/label= Table 3

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:46 :

CACCAGAGCAAACCGTCCAC

(2) INFORMATION FOR SEQ ID NO:47 :

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 20 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: double
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: primer

(ix) FEATURE:

- (A) NAME/KEY: misc\_feature
- (D) OTHER INFORMATION:/label= Table 3

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:47 :

ACAGCCCAGACTCCCATTCC

(2) INFORMATION FOR SEQ ID NO:48 :

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 20 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: double
  - (D) TOPOLOGY: linear



(ii) MOLECULE TYPE: primer

(ix) FEATURE:

(A) NAME/KEY: misc\_feature

(D) OTHER INFORMATION:/label= Table 3

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:48 :

TTCTCTTCTCCCTTCACCCCT

(2) INFORMATION FOR SEQ ID NO:49 :

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 22 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: double

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: primer

(ix) FEATURE:

(A) NAME/KEY: misc\_feature

(D) OTHER INFORMATION:/label= Table 3

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:49 :

ACACACTTCATGCTCTCTACCC

(2) INFORMATION FOR SEQ ID NO:50:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 20 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: double

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: primer

(ix) FEATURE:

(A) NAME/KEY: misc\_feature

(D) OTHER INFORMATION:/label= Table 3

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:50 :

CCGCCTATTCCCTTCCTCTT

(2) INFORMATION FOR SEQ ID NO:51 :

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 20 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: double

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: primer

(ix) FEATURE:

(A) NAME/KEY: misc\_feature

(D) OTHER INFORMATION:/label= Table 3

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:51 :

GACAAACTCCTGGGAAGCCT

(2) INFORMATION FOR SEQ ID NO:52 :

- (i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 20 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: double  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: primer

- (ix) FEATURE:  
(A) NAME/KEY: misc\_feature  
(D) OTHER INFORMATION:/label= Table 3

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:52 :

ACCTCTGACCCCTGTGAACC

(2) INFORMATION FOR SEQ ID NO:53:

- (i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 20 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: double  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: primer

- (ix) FEATURE:  
(A) NAME/KEY: misc\_feature  
(D) OTHER INFORMATION:/label= Table 3

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:53 :

TGTGGATTTGTGTGCTACGC

(2) INFORMATION FOR SEQ ID NO:54 :

- (i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 21 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: double  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: primer

- (ix) FEATURE:  
(A) NAME/KEY: misc\_feature  
(D) OTHER INFORMATION:/label= Table 3

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:54 :

CATAAATAGCACCGACAGGGA

(2) INFORMATION FOR SEQ ID NO:55 :

- (i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 20 base pairs

Sub  
C1

(B) TYPE: nucleic acid  
(C) STRANDEDNESS: double  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: primer

(ix) FEATURE:

(A) NAME/KEY: misc feature  
(D) OTHER INFORMATION:/label= Table 3

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:55 :

GGGATGGAGAAGAGTGAGGA

(2) INFORMATION FOR SEQ ID NO:56 :

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 20 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: double  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: primer

(ix) FEATURE:

(A) NAME/KEY: misc feature  
(D) OTHER INFORMATION:/label= Table 3

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:56 :

TCCTCACTCTTCTCCATCCC

(2) INFORMATION FOR SEQ ID NO:57 :

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 19 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: double  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: primer

(ix) FEATURE:

(A) NAME/KEY: misc feature  
(D) OTHER INFORMATION:/label= Table 3

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:57 :

ACCCTGTATGTTGCCTTGG

(2) INFORMATION FOR SEQ ID NO:58 :

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 20 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: double  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: primer

(ix) FEATURE:

(A) NAME/KEY: misc\_feature  
(D) OTHER INFORMATION:/label= Table 3

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:58 :

GGGGATTTTGCTGTGTGCTG

(2) INFORMATION FOR SEQ ID NO:59 :

(i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 20 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: double  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: primer

(ix) FEATURE:  
(A) NAME/KEY: misc\_feature  
(D) OTHER INFORMATION:/label= Table 3

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:59 :

ATTCCTGCTCCCACCGTCTC

(2) INFORMATION FOR SEQ ID NO:60:

(i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 20 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: double  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: primer

(ix) FEATURE:  
(A) NAME/KEY: misc\_feature  
(D) OTHER INFORMATION:/label= Table 3

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:60 :

CACAGAGTGTCCGAGAGGCA

(2) INFORMATION FOR SEQ ID NO:61 :

(i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 22 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: double  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: primer

(ix) FEATURE:  
(A) NAME/KEY: misc\_feature  
(D) OTHER INFORMATION:/label= Table 3

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:61 :

GGAGATTATCAGGTGAGATGCC

(2) INFORMATION FOR SEQ ID NO:62 :

- (i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 21 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: double  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: primer

- (ix) FEATURE:  
(A) NAME/KEY: misc\_feature  
(D) OTHER INFORMATION:/label= Table 3

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:62 :

CAGAGTGTCCGAGAGGCAGGG

(2) INFORMATION FOR SEQ ID NO:63 :

- (i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 20 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: double  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: primer

- (ix) FEATURE:  
(A) NAME/KEY: misc\_feature  
(D) OTHER INFORMATION:/label= Table 3

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:63 :

CGTTGACCCCTCCACCTTGA

(2) INFORMATION FOR SEQ ID NO:64 :

- (i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 20 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: double  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: primer

- (ix) FEATURE:  
(A) NAME/KEY: misc\_feature  
(D) OTHER INFORMATION:/label= Table 3

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:64 :

GGGAAAAACATGCACCTTCTT

(2) INFORMATION FOR SEQ ID NO:65:

- (i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 20 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: double  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: primer

(ix) FEATURE:

(A) NAME/KEY: misc\_feature

(D) OTHER INFORMATION:/label= Table 3

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:65 :

TAGGGGGTAAAAATGGAGGAG

(2) INFORMATION FOR SEQ ID NO:66 :

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 20 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: double

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: primer

(ix) FEATURE:

(A) NAME/KEY: misc\_feature

(D) OTHER INFORMATION:/label= Table 3

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:66 :

ACTAACTCAGTGGAATAGGG

(2) INFORMATION FOR SEQ ID NO:67 :

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 20 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: double

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: primer

(ix) FEATURE:

(A) NAME/KEY: misc\_feature

(D) OTHER INFORMATION:/label= Table 3

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:67 :

GGAGCTAGGATAGCTCAAT

□